

Eigenvalue confinement and spectral gap for random simplicial complexes

Antti Knowles*

Ron Rosenthal†

September 8, 2015

Abstract

We consider the adjacency operator of the Linial-Meshulam model for random simplicial complexes on n vertices, where each d -cell is added independently with probability p to the complete $(d-1)$ -skeleton. Under the assumption $np(1-p) \gg \log^4 n$, we prove that the spectral gap between the $\binom{n-1}{d}$ smallest eigenvalues and the remaining $\binom{n-1}{d-1}$ eigenvalues is $np - 2\sqrt{dnp(1-p)}(1+o(1))$ with high probability. This estimate follows from a more general result on eigenvalue confinement. In addition, we prove that the global distribution of the eigenvalues is asymptotically given by the semicircle law. The main ingredient of the proof is a Füredi-Komlós-type argument for random simplicial complexes, which may be regarded as sparse random matrix models with dependent entries.

1 Introduction

The Erdős-Rényi graph [ER59, ER61] $G(n, p)$ is a random graph on n vertices, where each edge is added independently with probability p . The spectrum of its adjacency matrix has been extensively studied [FK81, FKS89, FO05, CO07, HKP12, EKYY12, EKYY13]. Generally, the spectrum of the adjacency matrix of a graph encodes many important properties of the graph, in particular relating to its connectivity and expansion properties [AM85, Alo86, Nil91, Chu97, HLW06, KS06, CRS10]. From the point of view of random matrix theory, the adjacency matrix of $G(n, p)$ is a symmetric sparse random matrix with independent upper-triangular entries.

In this paper we study the spectra of random simplicial complexes. We consider the *Linial-Meshulam model* [LM06], which is high-dimensional generalization of the Erdős-Rényi model. Given $n, d \in \mathbb{N}$ and $p \in [0, 1]$, the Linial-Meshulam model $X \equiv X(d, n, p)$ is a random d -dimensional complex on n vertices with a complete $(d-1)$ -skeleton in which each d -cell is added independently with probability p . For $d = 1$, the Linial-Meshulam model reduces to the Erdős-Rényi model $G(n, p)$. Following its introduction in [LM06], the Linial-Meshulam has been extensively studied in [MW09, Koz10, BHK11, Wag11, CCFK12, HKP12, HJ13, HKP13, ALM13, GW14, LP14]. The notion of adjacency matrix has a natural extension to simplicial complexes, whereby the adjacency operator of a complex X is a self-adjoint operator that encodes the information whether two $(d-1)$ -cells belong to a common d -cell or not.

As for graphs, the spectrum of the adjacency operator, in particular its spectral gap, determines a notion of spectral expansion. There has recently been considerable interest in *high-dimensional expanders*, namely analogs of expander graphs in the context of general simplicial complexes. Unlike

*Partially supported by Swiss National Science Foundation grant 144662.

†Partially supported by an ETH fellowship.

the graph case $d = 1$, where various different notions of expansion are closely related, in the high-dimensional case $d > 1$ the several notions of expansion that have been proposed in the literature are in general far from being equivalent, and the relationship between them is still poorly understood. Notions of expansion for $d > 1$ that have been proposed include the aforementioned spectral expansion [Gar73, GW14, GP14, Opp14], combinatorial expansion [PRT15, Par13, Gol13, GS14, CMRT14], geometric and topological expansion [Gro10, FGL⁺12, MW14, DKW15, Evr15], \mathbb{F}_2 -coboundary expansion [LM06, MW09, Gro10, DK12, SKM14, GW14, LM15] and Ramanujan complexes [CSŻ03, Li04, LSV05, GP14, EGL15, KKL14]. In addition, the adjacency matrix can be interpreted as a generator of a stochastic process which is a high-dimensional analog of simple random walks, see [PR12, MS13, Ros14] for more details. A recent survey can be found in [Lub14].

From the point of view of random matrix theory, the adjacency matrix of the Linial-Meshulam model is a sparse self-adjoint random matrix. The entries of this matrix are independent (up to the self-adjointness constraint) if and only if $d = 1$. Indeed, the independent random variables are associated with d -cells, while entries of the adjacency matrix are associated with *pairs of* $(d - 1)$ -cells; these two notions coincide if and only if $d = 1$. The algebraic relationship between the independent random variables and the matrix entries is governed by the simplicial structure.

We now give an informal summary of our results. Throughout the following we use the abbreviation

$$q := p(1 - p). \quad (1.1)$$

Under the assumption $nq \gg \log^4 n$ and fixed d , we prove that the $N := \binom{n}{d} = \binom{n-1}{d} + \binom{n-1}{d-1}$ eigenvalues of the adjacency operator are with high probability confined to two separate intervals: denoting by $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_N$ the eigenvalues of the adjacency operator, we have with high probability

$$\lambda_i \in \begin{cases} \sqrt{dnq} [-2 - o(1), 2 + o(1)] & \text{if } i \leq \binom{n-1}{d} \\ np + [-7d, 7d] & \text{if } i > \binom{n-1}{d} \end{cases}. \quad (1.2)$$

The first estimate is optimal, while the second one is not (in fact the constant 7 may be easily improved). See Figure 1.1 for an illustration. As an immediate corollary, we obtain the spectral gap

$$\lambda_{\binom{n-1}{d}+1} - \lambda_{\binom{n-1}{d}} = np - 2\sqrt{dnq} (1 + o(1)). \quad (1.3)$$

We refer to Theorem 2.1 and Corollary 2.4 below for the precise statements.

Previously, a related confinement result for the eigenvalues of the Linial-Meshulam model was established in [GW14, Theorem 2], where the authors assume that $nq \gg \log n$ and establish (1.2) with the larger intervals $\sqrt{nq}[-C, C]$ for $i \leq \binom{n-1}{d}$ and $np + \sqrt{nq}[-C, C]$ for $i > \binom{n-1}{d}$, where C is some positive constant. This implies a spectral gap (1.3) equal to $np + O(\sqrt{nq})$. Hence, at the cost of the stronger assumption $nq \gg \log^4 n$ instead of $nq \gg \log n$, we improve the confinement of [GW14] by obtaining the optimal constant for the size of the interval in the first case of (1.2), and by improving the size of the interval in the second case of (1.2) by a factor $(nq)^{-1/2}$. Together, these improvements allow us to identify the subleading term in the spectral gap (1.3).

Finally, we prove that when $nq \rightarrow \infty$ the empirical spectral measure is asymptotically given by Wigner's semicircle law [Wig58] on the interval $\sqrt{dnq}[-2, 2]$. In particular, this implies the optimality of the first bound of (1.2). This result extends the well-known semicircle law for $G(n, p)$ to high-dimensional simplicial complexes.

We conclude this section with a few words about the proof. The main ingredient of the proof is a Füredi-Komlós-type argument for random simplicial complexes. The Füredi-Komlós argument involves estimating the expectation of the trace of a very high power of the centered adjacency matrix, by encoding the many resulting terms using walks on graphs. A fundamental observation in the original Füredi-Komlós argument is that, owing to the vanishing expectation of the entries of the centered

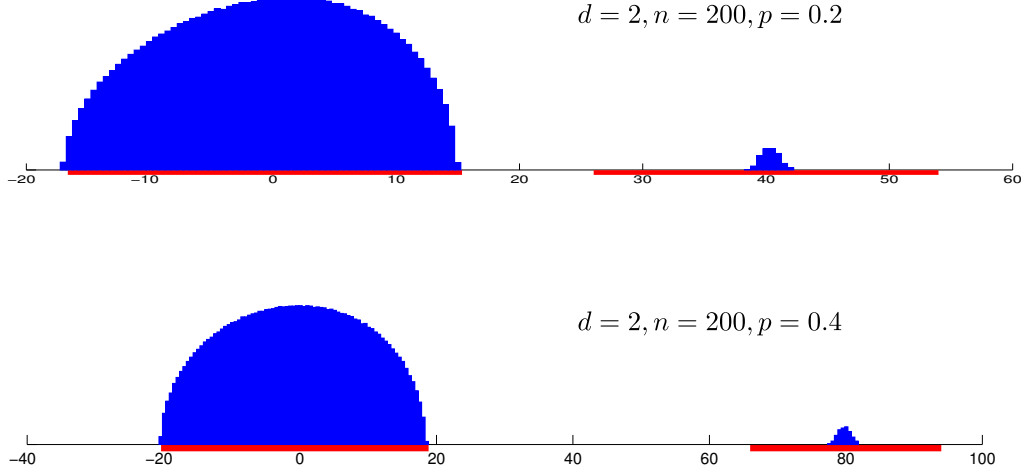


Figure 1.1: A histogram of the eigenvalue distribution of A for the values $(d, n, p) = (2, 200, 0.2)$ (top) and $(d, n, p) = (2, 200, 0.4)$ (bottom). The two intervals on the right-hand side of (1.2) are indicated using red regions.

adjacency matrix, each edge of the graph must be crossed at least twice, leading to a reduction in the number of admissible walks. For $d > 1$, this is no longer true because of the dependencies among the matrix entries. Obtaining sharp enough upper bounds on the number of admissible walks represents the main work in our proof. As it turns out, for $d > 1$ the mechanism behind the reduction in the number of admissible walks is more subtle, and, unlike for $d = 1$, nonlocal. We refer to Section 5 for a more detailed discussion on how to estimate the number of admissible paths. Finally, in Section 6, in order to prove the second case of (1.2), we need to modify the argument described above to obtain smaller bounds, by a factor $(nq)^{-1/2}$, for the restriction of the adjacency matrix to an explicit $\binom{n-1}{d-1}$ -dimensional subspace (denoted by $\text{im } P$ in Section 6). This estimate is obtained by a twist of the argument developed in Section 5. Our main results then follow easily from these estimates, combined with eigenvalue interlacing bounds and second order perturbation theory (see Sections 4 and 6).

2 Results

Let X be a finite simplicial complex with vertex set V of size n . This means that X is a finite collection of subsets of V , called *cells*, which is closed under taking subsets, i.e., if $\tau \in X$ and $\sigma \subseteq \tau$, then $\sigma \in X$. The *dimension* of a cell σ is $|\sigma| - 1$, and X^j denotes the set of cells of dimension j , which we call *j-cells*. The dimension of X , which we denote by d , is the maximal dimension of a cell in it. We assume that the complex has a *complete* $(d - 1)$ -*skeleton*, by which we mean that X contains all subsets of V of size at most d .

For $j \geq 1$, every j -cell $\sigma = \{\sigma^0, \dots, \sigma^j\} \in X^j$ has two possible orientations, corresponding to the possible orderings of its vertices, up to an even permutation. We denote an oriented cell by square brackets, and a flip of orientation by an overline. For example, one orientation of $\sigma = \{x, y, z\}$ is $[x, y, z] = [y, z, x] = [z, x, y]$. The other orientation of σ is $\overline{[x, y, z]} = [y, x, z] = [x, z, y] = [z, y, x]$. We denote by X_{\pm}^j the set of oriented j -cells (so that $|X_{\pm}^j| = 2|X^j|$ for $j \geq 1$). Moreover, we set $X_{\pm}^0 = X^0 = V$.

For $j \geq 1$, the space of *j-forms on X*, denoted by $\Omega^j(X)$, is the vector space of skew-symmetric

functions on oriented j -cells:

$$\Omega^j \equiv \Omega^j(X) := \left\{ f : X_{\pm}^j \rightarrow \mathbb{R} \mid f(\bar{\sigma}) = -f(\sigma) \forall \sigma \in X_{\pm}^j \right\}.$$

In particular, Ω^1 is the space of flows on edges. We endow each Ω^j with the inner product

$$\langle f, g \rangle = \sum_{\sigma \in X^j} f(\sigma)g(\sigma).$$

Note that $f(\sigma)g(\sigma)$ is well-defined since its value is independent of the choice of orientation of the j -cell σ .

Next, define the boundary $\partial\sigma$ of the $(j+1)$ -cell $\sigma = \{\sigma^0, \dots, \sigma^{j+1}\} \in X^{j+1}$ as the set of j -cells $\{\sigma^0, \dots, \sigma^{i-1}, \sigma^{i+1}, \dots, \sigma^j\}$ for $0 \leq i \leq j+1$. An oriented $(j+1)$ -cell $[\sigma^0, \dots, \sigma^{j+1}] \in X_{\pm}^{j+1}$ induces orientations on the j -cells in its boundary, as follows: the cell $\{\sigma^0, \dots, \sigma^{i-1}, \sigma^{i+1}, \dots, \sigma^j\}$ is oriented as $(-1)^i[\sigma^0, \dots, \sigma^{i-1}, \sigma^{i+1}, \dots, \sigma^j]$, where we use the notation $(-1)\tau := \bar{\tau}$.

The following neighboring relation for oriented cells was introduced in [PR12]: for $\sigma, \sigma' \in X_{\pm}^{d-1}$ we denote $\sigma' \sim \sigma$ (or $\sigma \overset{X}{\sim} \sigma'$) if there exists an oriented d -cell $\tau \in X_{\pm}^d$ such that both σ and $\bar{\sigma}'$ are in the boundary of τ as oriented cells (see Figure 2.1 for an illustration in the case $d = 2$).



Figure 2.1: Left: an oriented 2-cell and the orientation it induces on its boundary. Right: an oriented 1-cell in a 2-cell together with its two oriented neighboring 1-cells.

The adjacency operator $A = A_X$ of a complex X on the space $\Omega^{d-1}(X)$ is defined by

$$Af(\sigma) := \sum_{\sigma' \sim \sigma} f(\sigma'), \quad \forall f \in \Omega^{d-1}(X), \sigma \in X_{\pm}^{d-1}. \quad (2.1)$$

This definition is a rather direct way to introduce the adjacency operator. An equivalent, and more conceptual, definition is $A = D - \Delta^+$, where D is the degree operator of $(d-1)$ -cells (the *degree* of a j -cell is the number of $(j+1)$ -cells which contain it) and Δ^+ is the upper Laplacian originating in the work of Eckmann [Eck45]. More on the connection between A, Δ^+ and the homology and cohomology of the complex can be found in [Hat02, GW14]. See also [PRT15, MS13, Ros14] for the connection to random walks on simplicial complexes.

We denote by $K \equiv K(d, n)$ the complete d -complex on the n vertices in V and by $\mathbb{A} := A_K$ its adjacency operator.

Coming back to the Linial-Meshulam model, for $n, d \in \mathbb{N}$ satisfying $n \geq d+1$, and $p = p(n) \in [0, 1]$, the Linial-Meshulam complex $X = X(d, n, p)$ is a random d -dimensional complex on n vertices with a complete $(d-1)$ -skeleton in which each d -cell of K is added to X independently with probability p . This in particular implies that $X^j = K^j$ for every $0 \leq j \leq d-1$ and thus also $\Omega^j(X) = \Omega^j(K)$ for $0 \leq j \leq d-1$. We denote by $\mathbb{P} \equiv \mathbb{P}_n$ and $\mathbb{E} \equiv \mathbb{E}_n$ the probability measure and expectation with respect to the law of X . Note that for $d = 1$ this is exactly the Erdős-Rényi model $G(n, p)$.

Throughout this paper we fix $d \geq 2$. All of our results also hold for the case $d = 1$ which was already extensively studied in the literature; see [Wig58, FK81, FO05, CO07, HKP12] and the references

therein. The adjacency operator of X , $A = A_X$, is a self adjoint operator on the $N \equiv |X^{d-1}| = \binom{n}{d}$ -dimensional space $\Omega^{d-1}(X)$.

We now state our main results.

Theorem 2.1 (Eigenvalue confinement). *For every $d \geq 2$ there exists a positive constant $C > 0$ depending only on d such that the following holds with probability at least $1 - n^{-D}$ for all $D > 0$.*

- (1) *For every $\xi > 0$, if $nq \geq \frac{C(1+D)^4}{\min\{\xi^6, 1\}} \log^4 n$ then the $\binom{n-1}{d}$ smallest eigenvalues of A lie in the interval $\sqrt{dnq}[-2 - \xi, 2 + \xi]$.*
- (2) *If $nq \geq C(1+D)^4 \log^4 n$ and $q \log^6 n \leq \frac{1}{C(1+D)^6}$, then the $\binom{n-1}{d-1}$ largest eigenvalues of A lie in the interval $np + [-7d, 7d]$.*

Remark 2.2. In Theorem 6.1 below, we give a slightly more precise version of Theorem 2.1. For instance, we can replace (2) with the following statement.

(2') If $nq \geq C(1+D)^4 \log^4 n$ then the $\binom{n-1}{d-1}$ largest eigenvalues of A lie in the interval

$$np + (6d + O_D(\sqrt{q} \log^3 n))[-1, 1]. \quad (2.2)$$

From Theorem 2.1 and Remark 2.2, we immediately get the following result.

Corollary 2.3 (Spectral gap). *For every $d \geq 2$ there exists a positive constant $C > 0$ depending only on d such that for all $\xi > 0$ and $D > 0$ satisfying $nq \geq \frac{C(1+D)^4}{\min\{\xi^6, 1\}} \log^4 n$ we have*

$$\lambda_{\binom{n-1}{d}+1} - \lambda_{\binom{n-1}{d}} = np - 2\sqrt{dnq} (1 + O_D(\xi))$$

with probability at least $1 - n^{-D}$.

For the following results, where $n \rightarrow \infty$ in probability space, we couple all random complexes $(X(d, n, p))_{n \in \mathbb{N}}$ in the standard fashion. Namely, we work on the probability space generated by the family of i.i.d. random variables $(\chi_\tau)_\tau$ indexed by all d -cells τ on the infinite vertex set $\{1, 2, 3, \dots\}$, where χ_τ is a Bernoulli- p random variable. Thus, for any $n \geq 1$ and a d -cell $\tau \subset \{1, \dots, n\}$, we set τ to belong to $X(d, n, p)$ if and only if $\chi_\tau = 1$.

Corollary 2.4. *Fix $d \geq 2$ and an open interval I containing $[-2, 2]$. If $\lim_{n \rightarrow \infty} \frac{nq}{\log^4 n} = \infty$, then, almost surely, the $\binom{n-1}{d}$ smallest eigenvalues of A lie in $\sqrt{dnq} I$ for all but finitely many values of n .*

An analogous result holds for the largest $\binom{n-1}{d-1}$ eigenvalues, whose precise statement we omit.

Our final result is about the eigenvalue distribution of A . For a self-adjoint operator B operating on an N -dimensional space, denote by $(\lambda_i(B))_{i=1}^N$ its eigenvalues in non-decreasing order and by

$$L_B = \frac{1}{N} \sum_{i=1}^N \delta_{\lambda_i(B)} \quad (2.3)$$

the empirical spectral measure of B .

Theorem 2.5. *If $d \geq 2$ and $\lim_{n \rightarrow \infty} nq = \infty$, then, almost surely,*

$$L_{(dnq)^{-1/2}A} \xrightarrow{\mathcal{D}} \nu, \quad (2.4)$$

where $\nu(dx) := \frac{\sqrt{(4-x^2)_+}}{2\pi} dx$ is the semicircle distribution.

Remark 2.6. It follows from Theorem 2.5 that the constant 2 in the interval $\sqrt{dnq}[-2 - \xi, 2 + \xi]$ from Theorem 2.1(1) is optimal. In fact, it is an easy corollary of Theorems 2.1 and 2.5 that if $\lim_{n \rightarrow \infty} \frac{nq}{\log^4 n} = \infty$ then $\lim_{n \rightarrow \infty} (dnq)^{-1/2} \lambda_{\binom{n-1}{d}} = 2$ and $\lim_{n \rightarrow \infty} (dnq)^{-1/2} \lambda_1 = -2$ almost surely. See also Remark 5.2 below.

Conventions. We use C to denote a generic large positive constant, which may depend on some fixed parameters and whose value may change from one expression to the next. If C depends on some parameter k , we sometimes emphasize this dependence by writing C_k instead of C . Moreover, for $f, g : \mathbb{N} \rightarrow \mathbb{R}$ we write $f(n) = O_k(g(n))$ to mean $f(n) \leq C_k g(n)$ for all $n \in \mathbb{N}$. Finally, we abbreviate $\llbracket n \rrbracket := \{1, \dots, n\}$. The letters $d, i, j, k, l, m, n, s, N$ are always used to denote an element in $\mathbb{N} = \{0, 1, 2, \dots\}$.

From now on, we consistently use σ for (oriented or non-oriented) $(d-1)$ -cells, and τ for (oriented or non-oriented) d -cells.

3 Semicircle law for $A - \mathbb{E}[A]$

In this section we prove the semicircle law for the centered and normalized adjacency matrix

$$H := \frac{1}{\sqrt{nq}}(A - \mathbb{E}[A]). \quad (3.1)$$

More precisely, we prove that the empirical spectral measure L_H converges to the rescaled distribution $\nu_d(dx) := \frac{\sqrt{(4d-x^2)_+}}{2\pi d} dx$.

Theorem 3.1. *Fix $d \geq 2$. If $\lim_{n \rightarrow \infty} nq = \infty$, then as $n \rightarrow \infty$ almost surely $L_H \xrightarrow{\mathcal{D}} \nu_d$.*

Using Theorem 3.1, we shall conclude the proof of Theorem 2.5 in Section 4 below. The proof of Theorem 3.1 is based on the moment method (see e.g. the presentation of [AGZ10, Section 2.1] for the classical case of independent matrix entries), and is the subject of the rest of this section. Along the proof, we also record several definitions and notions that will be used for the proof of Theorem 2.1 in Sections 5–6.

By a standard truncation argument and the Borel-Cantelli lemma, Theorem 3.1 follows from the two next lemmas (see [AGZ10, Section 2.1.2] for details).

Lemma 3.2. *If $\lim_{n \rightarrow \infty} nq = \infty$, then for every fixed $k \in \mathbb{N}$*

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\int_{\mathbb{R}} x^k L_H(dx) \right] = \int_{\mathbb{R}} x^k \nu_d(dx) = \begin{cases} 0 & \text{if } k \text{ is odd} \\ d^{k/2} \mathcal{C}_{k/2} & \text{if } k \text{ is even} \end{cases},$$

where $\mathcal{C}_k := \frac{1}{k+1} \binom{2k}{k}$ is the k -th Catalan number.

Lemma 3.3. *For every fixed $k \in \mathbb{N}$ we have*

$$\text{Var} \left(\int_{\mathbb{R}} x^k L_H(dx) \right) \leq O_k \left(\frac{1}{n^d(nq)} \right). \quad (3.2)$$

The rest of this section is devoted to the proof of Lemma 3.2. The proof of Lemma 3.3 is a standard adaptation of the ideas of the proof of Lemma 3.2, and we omit its details. We begin with definitions that we use throughout the remainder of the paper.

First, we need an explicit matrix representation of the adjacency operator. We fix an arbitrary choice of orientation of the $(d-1)$ -cells $K_+^{d-1} \subset K_{\pm}^{d-1}$, which in turn induces a choice of orientation on the $(d-1)$ -cells of X since $X^{d-1} = K^{d-1}$. Note that there is a natural bijection between K_+^{d-1} and K^{d-1} , and hence also between X_+^{d-1} and X^{d-1} . From now on, by a slight abuse of notation, using this bijection we often write $\sigma_1 \cup \sigma_2$ and $\sigma_1 \cap \sigma_2$ for oriented $(d-1)$ -cells σ_1, σ_2 to denote the union and intersection of the corresponding unoriented cells.

The orientation $K_+^{d-1} \subset K_\pm^{d-1}$ gives rise to an associated orthonormal basis $(\mathbb{1}_\sigma)_{\sigma \in X_+^{d-1}}$ of Ω^{d-1} , defined by

$$\mathbb{1}_\sigma(\sigma') = \begin{cases} 1 & \text{if } \sigma' = \sigma \\ -1 & \text{if } \sigma' = \bar{\sigma} \\ 0 & \text{otherwise} \end{cases}.$$

Then the adjacency matrix is an $X_+^{d-1} \times X_+^{d-1}$ matrix $(A_{\sigma\sigma'})$ with entries $A_{\sigma\sigma'} := \langle \mathbb{1}_\sigma, A\mathbb{1}_{\sigma'} \rangle$. Explicitly,

$$A_{\sigma,\sigma'} = \begin{cases} 1 & \text{if } \sigma \stackrel{\sim}{\sim} \sigma' \\ -1 & \text{if } \sigma \stackrel{\sim}{\sim} \bar{\sigma'} \\ 0 & \text{otherwise} \end{cases}. \quad (3.3)$$

In particular,

$$\mathbb{E}[A] = p\mathbb{A}, \quad (3.4)$$

where we recall that \mathbb{A} is the adjacency operator of the complete d -complex K .

Next, we introduce the basic definitions underlying the proof of Lemma 3.2. They are illustrated in Figure 3.1.

Definition 3.4 (Words). An (n, d) -letter σ (or shortly a *letter*) is an element of X_+^{d-1} . An (n, d) -word w (or shortly a *word*) is a finite sequence $\sigma_1 \dots \sigma_k$ of letters at least one letter long such that $\sigma_i \cup \sigma_{i+1}$ is a d -cell in K for every $1 \leq i \leq k-1$. The *length* of the word $\sigma_1 \dots \sigma_k$ is defined to be k . A word is called *closed* if its first and last letters are the same. Two words $w = \sigma_1 \dots \sigma_k$ and $w' = \sigma'_1 \dots \sigma'_k$ are called *equivalent* if there exists a permutation π on $X^0 = V$ such that $\pi(\sigma_i) = \sigma'_i$ for every $1 \leq i \leq k$, where for $\sigma = [\sigma^0, \dots, \sigma^{d-1}] \in X_\pm^{d-1}$ we write $\pi(\sigma) = [\pi(\sigma^0), \dots, \pi(\sigma^{d-1})]$.

Definition 3.5 (Support of a word). For a word $w = \sigma_1 \dots \sigma_k$ we define its *support* by $\text{supp}_0(w) = \sigma_1 \cup \sigma_2 \cup \dots \cup \sigma_k$ and its *d -cell support* by $\text{supp}_d(w) = \{\sigma_i \cup \sigma_{i+1} : 1 \leq i \leq k-1\}$.

Definition 3.6 (Graph of a word). Given a word $w = \sigma_1 \dots \sigma_k$ we define $G_w = (V_w, E_w)$ to be the graph with vertex set $V_w = \{\sigma_i : 1 \leq i \leq k\}$ and edge set $E_w = \{\{\sigma_i, \sigma_{i+1}\} : 1 \leq i \leq k-1\}$. To avoid confusion, from this point on we consistently use the words vertex and edge to describe elements of V_w and E_w respectively; elements in X^0 and X^1 are always referred to as 0-cells and 1-cells.

The graph G_w comes with a path, given by the word w , that goes through all of its vertices and edges. We call each step along the path, i.e. $\sigma_i \sigma_{i+1}$ for some $1 \leq i \leq k-1$, a *crossing* of the edge $\{\sigma_i, \sigma_{i+1}\}$ and a *crossing* of the d -cell $\sigma_i \cup \sigma_{i+1}$. The path gives a natural ordering of the vertices and edges of G_w , and of the d -cells in $\text{supp}_d(w)$ according to the order of their first crossing along w . Finally, for an edge $e \in E_w$ define $N_w(e)$ to be the number of times the edge e is crossed along the path generated by w in the graph G_w . Given an edge $e \in E_w$ and $1 \leq i \leq N_w(e)$, the *i -th crossing time of the edge e* is given by the unique $1 \leq j \leq k-1$ such that $\sigma_j \sigma_{j+1}$ is the i -th crossing of the edge e along the path w .

The edges of a graph G_w , associated with a word $w = \sigma_1 \dots \sigma_k$, can be divided into different classes according to the d -cell generated by the two $(d-1)$ -cells in its endpoints. This is done as follows.

Definition 3.7. For a d -cell τ let $E_w(\tau) = \{\{\sigma, \sigma'\} \in E_w : \sigma \cup \sigma' = \tau\}$ and define (with a slight abuse of notation)

$$N_w(\tau) = \sum_{e \in E_w(\tau)} N_w(e)$$

to be the total number of times the d -cell τ is crossed along the path generated by the word w . As for edges of the graph, given $\tau \in X^d$ and $1 \leq i \leq N_w(\tau)$ we define the *i -th crossing time of a d -cell τ* to be the unique $1 \leq j \leq k-1$ such that $\sigma_j \sigma_{j+1}$ is the i -th crossing of the d -cell τ along the path w .

For $\sigma, \sigma' \in X_+^{d-1}$ the condition $\sigma \cup \sigma' \in K^d$ can mean one of two things, either $\sigma \stackrel{K}{\sim} \sigma'$ or $\sigma \stackrel{K}{\sim} \overline{\sigma'}$. For future use we denote the set of non-neighboring edges whose union is τ by

$$\widehat{E}_w(\tau) = \{\{\sigma, \sigma'\} \in E_w(\tau) : \sigma \stackrel{K}{\sim} \overline{\sigma'}\},$$

so that $E_w(\tau) \setminus \widehat{E}_w(\tau) = \{\{\sigma, \sigma'\} \in E_w(\tau) : \sigma \stackrel{K}{\sim} \sigma'\}$.

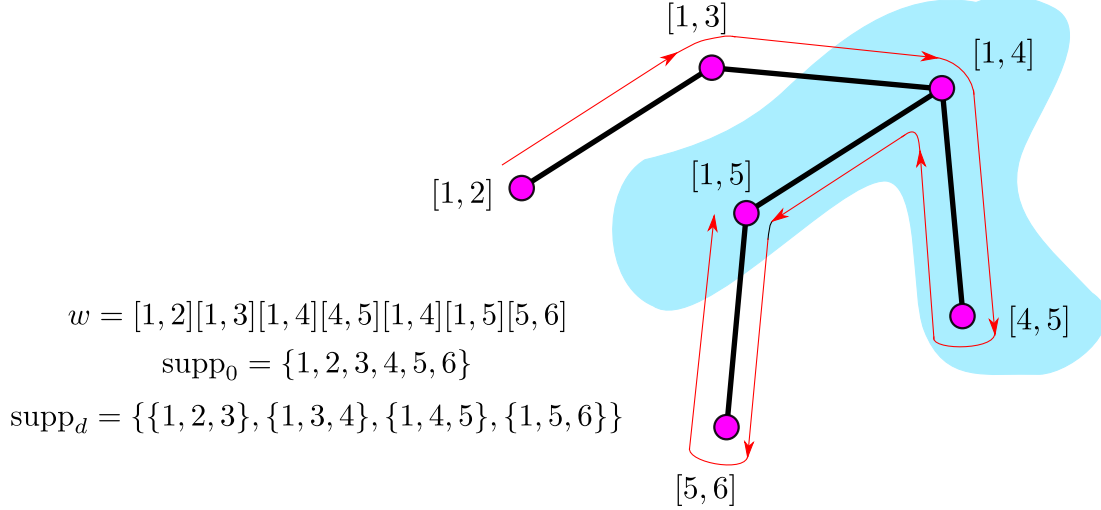


Figure 3.1: We illustrate Definitions 3.4-3.7 for the word $w = [1,2][1,3][1,4][4,5][1,4][1,5][5,6]$ (note that w is indeed a word since the union of any two consecutive 1-cells is a 2-cell). The 0-cell support $\text{supp}_0(w)$ and d -cell support $\text{supp}_d(w)$ (for $d = 2$) of the word w are given in the figure and the associated graph is drawn. The path induced by w on the graph G_w is illustrated by the red path. Note that the edges $\{[1,4], [1,5]\}$ and $\{[1,4], [4,5]\}$ (see blue area) belong to the same 2-cell $\{1,4,5\}$. We therefore have that $N_w(\{1,4,5\}) = N_w(\{[1,4], [1,5]\}) + N_w(\{[1,4], [4,5]\}) = 2 + 1 = 3$. As for the non-neighboring edges, we have for example $\widehat{E}_w(\{1,4,5\}) = \{\{[1,4], [4,5]\}\}$. The first crossing time of the edge $\{[1,4], [4,5]\}$ is 3 and the third crossing time of the d -cell $\{1,4,5\}$ is 5.

We remark that for $d = 1$, there is no distinction between $N_w(\tau)$ and $N_w(e)$, since in that case the edges of the graph are naturally associated with 1-cells.

Abbreviate $B = A - \mathbb{E}[A] = A - p\mathbb{A}$. For a fixed $k \in \mathbb{N}$, we wish to understand the limiting behavior of $\mathbb{E} \left[\int_{\mathbb{R}} x^k L_H(dx) \right]$. Since for $k = 0$ we have $\mathbb{E} \left[\int_{\mathbb{R}} L_H(dx) \right] = 1$ we will assume that $k \geq 1$. The starting point of the moment method is the identity

$$\mathbb{E} \left[\int_{\mathbb{R}} x^k L_H(dx) \right] = \frac{1}{N(nq)^{k/2}} \sum_{\sigma_1, \dots, \sigma_k \in X_+^{d-1}} \mathbb{E}[B_{\sigma_1 \sigma_2} B_{\sigma_2 \sigma_3} \dots B_{\sigma_{k-1} \sigma_k} B_{\sigma_k \sigma_1}], \quad (3.5)$$

which follows easily from (2.3).

Note that each term in the sum can be associated with a string of letters $\sigma_1 \sigma_2 \dots \sigma_k \sigma_1$. Since $B_{\sigma, \sigma'} = 0$ whenever $\sigma \cup \sigma' \notin X^d$ it follows that we can restrict the sum in (3.5) to the case where $\sigma_i \cup \sigma_{i+1} \in X^d$ for $1 \leq i \leq k-1$ and $\sigma_k \cup \sigma_1 \in X^d$. Consequently, the list of letters in the sum can be restricted to the set of closed words of length $k+1$. Using the independence structure of A for different d -cells and the definition of N_w we then have

$$\mathbb{E} \left[\int_{\mathbb{R}} x^k L_H(dx) \right] = \sum_{\substack{w \text{ a closed word} \\ \text{of length } k+1}} \frac{1}{N(nq)^{k/2}} \prod_{\tau \in X^d} \mathbb{E} \left[(\chi - p)^{N_w(\tau)} \right] \text{sgn}(w, \tau), \quad (3.6)$$

where χ is a Bernoulli random variable with parameter p and

$$\text{sgn}(w, \tau) = (-1)^{\sum_{e \in \widehat{E}_w(\tau)} N_w(e)} \quad (3.7)$$

is a sign that depends on the parity of the number of crossings between non-neighboring $(d-1)$ -cells in w . We write $T(w)$ for the contribution of the word w to the sum in (3.6), that is

$$T(w) = \frac{1}{N(nq)^{k/2}} \prod_{\tau \in X^d} \mathbb{E} \left[(\chi - p)^{N_w(\tau)} \right] \text{sgn}(w, \tau). \quad (3.8)$$

Since $\mathbb{E}[\chi - p] = 0$, it follows that $T(w) = 0$ unless $N_w(\tau) \neq 1$ for every $\tau \in X^d$. We can thus restrict the sum in (3.6) to words such that $N_w(\tau) \neq 1$ for every $\tau \in X^d$.

Recalling the definition of the equivalence relation for words, see Definition 3.4, we denote by $[w]$ the equivalence class of w and observe that, because the term $T(w)$ is invariant under permutations of X^0 , we have $T(w) = T(w')$ for every pair of words w, w' such that $[w] = [w']$.

Using the last observation and (3.8) we can rewrite (3.6) as

$$\mathbb{E} \left[\int_{\mathbb{R}} x^k L_H(dx) \right] = \sum_w |[w]| T(w), \quad (3.9)$$

where the sum is over a set of representatives for the equivalence classes of closed words of length $k+1$ such that $N_w(\tau) \neq 1$ for every $\tau \in X^d$.

Next, we distinguish between different equivalence classes according to the number of 0-cells in their support. Denote by $\mathcal{W}_s^k = \mathcal{W}_s^k(n, d)$ a set of representatives for the equivalence classes of closed words of length $k+1$ such that $N_w(\tau) \neq 1$ for every $\tau \in X^d$ and $|\text{supp}_0(w)| = s$. Note that this parameter is independent of the choice of the representative in the equivalence class. Using the fact that $\sum_{\tau \in X^d} N_w(\tau) = k$ and recalling that we only consider words such that $N_w(\tau) \neq 1$ for every $\tau \in X^d$, it follows that the number of d -cells such that $N_w(\tau) > 0$ is bounded by $\lfloor \frac{k}{2} \rfloor$.

Claim 3.8. $|\text{supp}_0(w)| \leq |\text{supp}_d(w)| + d$ for every word w .

Proof. The word w starts in a $(d-1)$ -cells which contains d distinct 0-cells. In order to obtain a new 0-cells in the i -th crossing, one must observe in the crossing $\sigma_i \sigma_{i+1}$ a new d -cell. Since there are $|\text{supp}_d(w)|$ distinct d -cells in w the result follows. \square

Consequently, $d \leq |\text{supp}_0(w)| \leq |\text{supp}_d(w)| + d \leq \lfloor \frac{k}{2} \rfloor + d$. Using (3.9) together with the last claim we thus conclude that

$$\begin{aligned} \mathbb{E} \left[\int_{\mathbb{R}} x^k L_H(dx) \right] &= \sum_{s=d}^{\lfloor \frac{k}{2} \rfloor + d} \sum_{w \in \mathcal{W}_s^k} |[w]| T(w) \\ &= \sum_{s=d}^{\lfloor \frac{k}{2} \rfloor + d} \frac{1}{N(nq)^{k/2}} \sum_{w \in \mathcal{W}_{t,s}^k} |[w]| \prod_{\tau \in X^d} \mathbb{E} \left[(\chi - p)^{N_w(\tau)} \right] \text{sgn}(w, \tau). \end{aligned} \quad (3.10)$$

We record the following simple claims, whose proofs are straightforward.

Claim 3.9. For every $w \in \mathcal{W}_s^k$, the number of elements in $[w]$ is $B_{n,s,d} := \frac{n(n-1)\cdots(n-s)}{d!}$ and in particular is bounded by n^s .

Claim 3.10. For every $k \geq 0$ and $d \leq s \leq \lfloor \frac{k}{2} \rfloor + d$ we have

$$|\mathcal{W}_s^k| \leq k^{dk}. \quad (3.11)$$

Also, observe that for $w \in \mathcal{W}_s^k$

$$\begin{aligned} \prod_{\tau \in X^d} \mathbb{E} \left[(\chi - p)^{N_w(\tau)} \right] &= \prod_{\substack{\tau \in X^d \\ N_w(\tau) \geq 2}} p(1-p) \left[(1-p)^{N_w(\tau)-1} + p^{N_w(\tau)-1} \right] \\ &\leq \prod_{\substack{\tau \in X^d \\ N_w(\tau) \geq 2}} p(1-p) = q^{|\text{supp}_d(w)|} \leq q^{|\text{supp}_0(w)|-d} = q^{s-d}, \end{aligned} \quad (3.12)$$

where for the last inequality we used Claim 3.8.

Combining (3.12), Claim 3.9 and Claim 3.10 we conclude that for odd k and large enough n

$$\begin{aligned} \left| \sum_{s=d}^{\lfloor \frac{k}{2} \rfloor + d} \frac{B_{n,s,d}}{N(nq)^{k/2}} \sum_{w \in \mathcal{W}_s^k} \prod_{\tau \in X^d} \mathbb{E} \left[(\chi - p)^{N_w(\tau)} \right] \text{sgn}(w, \tau) \right| &\leq \sum_{s=d}^{\lfloor \frac{k}{2} \rfloor + d} \frac{n^s q^{s-d}}{N(nq)^{k/2}} |\mathcal{W}_s^k| \\ &\leq \sum_{s=d}^{\lfloor \frac{k}{2} \rfloor + d} \frac{(nq)^s}{Nq^d (nq)^{k/2}} k^{dk} \stackrel{(1)}{\leq} \sum_{s=d}^{\lfloor \frac{k}{2} \rfloor + d} \frac{(nq)^{\lfloor \frac{k}{2} \rfloor + d}}{Nq^d (nq)^{k/2}} k^{dk} \leq \frac{k}{2} k^{dk} d! (nq)^{\lfloor \frac{k}{2} \rfloor - \frac{k}{2}} = O_{k,d} \left(\frac{1}{(nq)^{1/2}} \right), \end{aligned}$$

where for (1) we used that $\lim_{n \rightarrow \infty} nq = \infty$, and in particular $nq \geq 1$ for large enough n . This completes the proof for odd k .

Similarly, when k is even, one can separate the sum to $s = \frac{k}{2} + d$ which we denote by R_1 and the sum over $s < \frac{k}{2} + d$ which we denote by R_2 , and obtain

$$R_2 = O_{k,d} \left(\frac{1}{nq} \right).$$

Thus, for k even we have

$$\begin{aligned} \mathbb{E} \left[\int_{\mathbb{R}} x^k L_H(dx) \right] &= R_1 + O_{k,d} \left(\frac{1}{nq} \right) \\ &= \frac{B_{n, \frac{k}{2} + d, d}}{N(nq)^{k/2}} \sum_{w \in \mathcal{W}_{\frac{k}{2} + d}^k} \prod_{\tau \in X^d} \mathbb{E} \left[(\chi - p)^{N_w(\tau)} \right] \text{sgn}(w, \tau) + O_{k,d} \left(\frac{1}{nq} \right). \end{aligned}$$

The following Lemma contains the key estimates needed to complete the proof.

Lemma 3.11. *The following claims hold for even k and $w \in \mathcal{W}_{\frac{k}{2} + d}^k$.*

- (1) $N_w(\tau) \in \{0, 2\}$ for every d -cell τ .
- (2) $|E_w| = \frac{k}{2}$ and $N_w(e) = 2$ for every $e \in E_w$. In particular $\text{sgn}(w, \tau) = 1$ for every $\tau \in X^d$.
- (3) $|\mathcal{W}_{\frac{k}{2} + d}^k| = C_{\frac{k}{2}} d^{k/2}$.

Assuming Lemma 3.11, the proof of Lemma 3.2 is now complete. Indeed, using the first two equalities in (3.12) together with Lemma 3.11 yields for even k

$$\mathbb{E} \left[\int_{\mathbb{R}} x^k L_H(dx) \right] = \frac{n(n-1) \cdots (n - k/2 - d)}{d! N(nq)^{k/2}} C_{\frac{k}{2}} d^{k/2} q^{\frac{k}{2}} + O_{k,d} \left(\frac{1}{nq} \right), \quad (3.13)$$

which gives

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\int_{\mathbb{R}} x^k L_H(dx) \right] = C_{\frac{k}{2}} d^{k/2}.$$

What remains, therefore, is the proof of Lemma 3.11, which contains the main novelty of the proof. In the graph case, $d = 1$, the fact that each d -cell is crossed either twice or not at all implies the same for the edges of the graph G_w . For $d \geq 2$ this is no longer true, and it is not immediate that part (1) of Lemma 3.11 implies part (2), as each d -cell can be associated with up to $\binom{d+1}{2}$ different edges of G_w .

Proof of Lemma 3.11. Fix $w = \sigma_1 \sigma_2 \dots \sigma_k \sigma_{k+1} \in \mathcal{W}_{k/2+d}^k$ with $\sigma_{k+1} = \sigma_1$. We begin with the proof of (1). As observed before, the combination of the upper bound $|\text{supp}_d(w)| \leq \frac{k}{2}$ with Claim 3.8 implies that for $w \in \mathcal{W}_{k/2+d}^k$ we have $|\text{supp}_d(w)| = \frac{k}{2}$, which means that there are exactly $\frac{k}{2}$ distinct d -cells such that $N_w(\tau) \geq 2$ (recall that it is impossible to have $N_w(\tau) = 1$). Since w is also a closed word of length $k + 1$ we conclude that

$$\frac{k}{2} \cdot 2 \leq \sum_{\tau \in X^d} N_w(\tau) = \sum_{e \in E_w} N_w(e) = k$$

and therefore $N_w(\tau) = 2$ for each of the d -cells in $\text{supp}_d(w)$. This concludes the proof of (1).

Next, we prove (2). Let $\tau_1, \dots, \tau_{\frac{k}{2}}$ be the d -cells crossed along the path generated by w in the order of their appearance. Since the d -cells appear along the path generated by w , we must have that for every $1 \leq i \leq \frac{k}{2} - 1$ the d -cell τ_{i+1} is attached to one of the d -cells τ_1, \dots, τ_i along a joint $(d-1)$ -cell in their boundary, and in particular τ_{i+1} can add at most one new 0-cell to the 0-cells in $\tau_1 \cup \dots \cup \tau_i$. Noting that $\text{supp}_0(w) = \tau_1 \cup \dots \cup \tau_{\frac{k}{2}}$, and recalling that for $w \in \mathcal{W}_{k/2+d}^k$ we have $|\text{supp}_0(w)| = \frac{k}{2} + d$, it follows that for every $1 \leq i \leq \frac{k}{2}$ the d -cell τ_{i+1} contains exactly one 0-cell that does not belong to $\tau_1 \cup \dots \cup \tau_i$.

Next, we show that $|V_w \cap \partial\tau_i| = 2$ for every $1 \leq i \leq \frac{k}{2}$, and thus in particular that the number of edges in E_w is $\frac{k}{2}$, each of which is crossed precisely twice. The argument is separated into three cases, one of which is then further splitted into three subcases. Assume that there exists $\tau \in \text{supp}_d(w)$ such that $|V_w \cap \partial\tau| > 2$ and let $1 \leq i \leq \frac{k}{2}$ be the minimal index such that $|V_w \cap \partial\tau_i| > 2$. Assume further that $\sigma_l \sigma_{l+1}$ is the first crossing of τ_i and σ_m is the first appearance of a $(d-1)$ -cell in $\partial\tau_i$ which is not σ_l or σ_{l+1} . Note that it is impossible to have $m = l$ or $m = l + 1$.

- **Case 1:** $m < l$ (see Figure 3.2(a)). This case is impossible since it implies that the first crossing of τ_i does not add a new 0-cell to the ones in $\tau_1 \cup \dots \cup \tau_{i-1}$.
- **Case 2:** $m > l + 1$ and the first time the d -cell $\sigma_{m-1} \sigma_m$ appears along the path generated by w is in the crossing $\sigma_{m-1} \sigma_m$ (see Figure 3.2(b)). This is impossible, because it means that the d -cell $\tau_j = \sigma_{m-1} \sigma_m$ does not add a new 0-cell to the ones in $\tau_1 \cup \dots \cup \tau_{j-1}$ as it should.
- **Case 3:** $m > l + 1$ and the d -cell $\tau_j = \sigma_{m-1} \sigma_m$ already appeared along the path generated by w before the crossing $\sigma_{m-1} \sigma_m$.
 - **Case 3.1:** $j < i$. This is impossible, because it implies that τ_i is not the first d -cell with the property that $|V_w \cap \partial\tau| > 2$ in the list.
 - **Case 3.2:** $j = i$ (see Figure 3.2(c.2)). This is only possible if $m = l + 2$ (otherwise m is not the first appearance after $l + 1$ of a $(d-1)$ -cell from the boundary of τ_i). However, in this case one can define a new word $w' = \sigma_{l+2} \dots \sigma_{k+1} \sigma_2 \dots \sigma_l \sigma' \sigma_l$, where σ' is a $(d-1)$ -cell which is a neighbor of σ_l and contains a 0-cell that does not belong to $\text{supp}_0(w)$. The word w' has the same number of d -cells as the original word w , namely $|\text{supp}_d(w')| = \frac{k}{2}$, but has one more 0-cell than w , i.e., $|\text{supp}_0(w')| = \frac{k}{2} + d + 1$. This however contradicts Claim 3.8.
 - **Case 3.3:** $j > i$ (see Figure 3.2(c.3)). Denote by $\{\sigma_r, \sigma_{r+1}\}$ the first edge crossed in the d -cell τ_j with $l + 1 < r < m$. This gives yet another contradiction since the first appearance

of the d -cell τ_j did not add a new 0-cell to the ones in $\tau_1 \cup \dots \cup \tau_{j-1}$ as it should (we know that σ_{r+1} is obtained from σ_r by removing one vertex and necessarily adding a vertex which belongs to τ_i or otherwise σ_m does not belong to $\partial\tau_i$).

Since all cases leads to contradiction we conclude that $|V_w \cap \partial\tau_i| = 2$ for every $1 \leq i \leq \frac{k}{2}$ and in particular, there is exactly one edge associated with each d -cell, namely $|E_w(\tau)| = 1$ for every $\tau \in \text{supp}_d(w)$. This concludes the proof of (2).

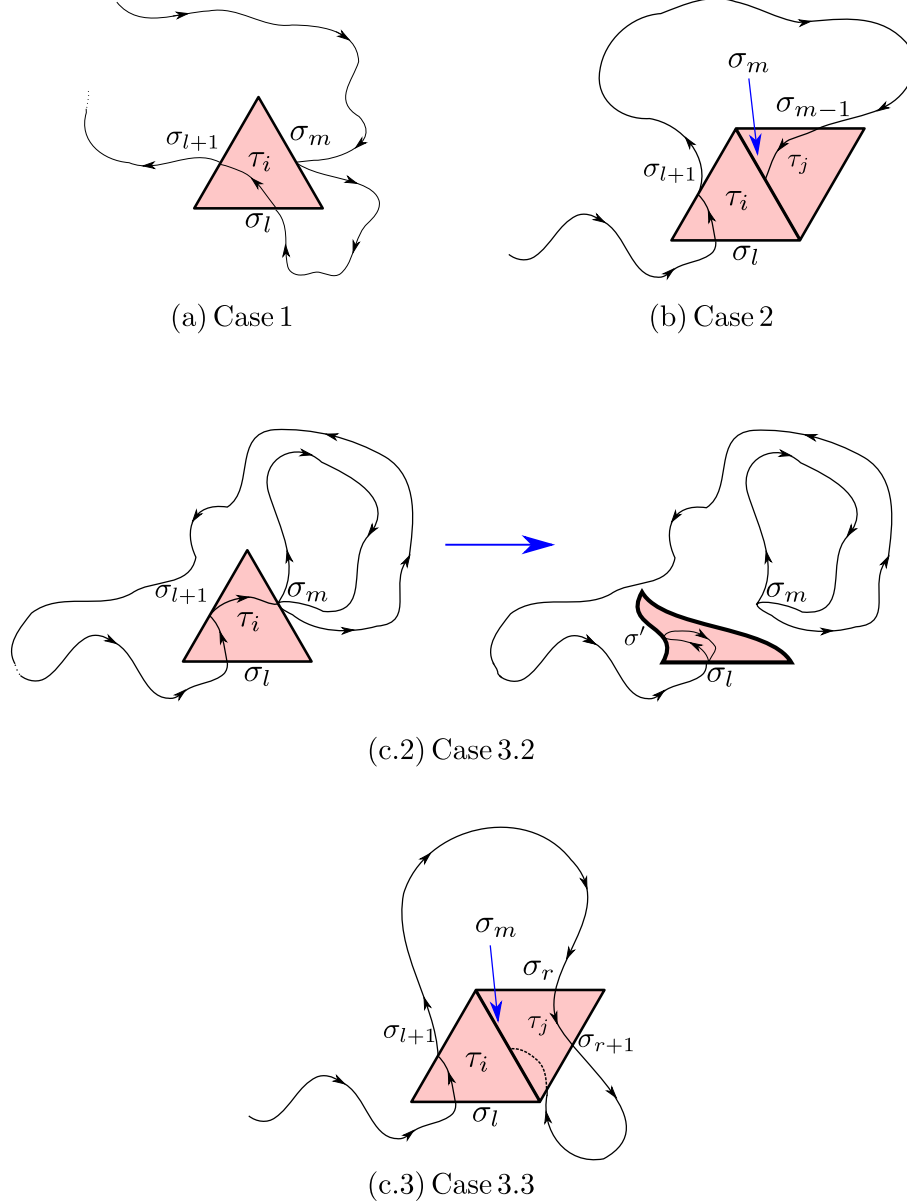


Figure 3.2: An illustration for the different cases of a d -cell τ with $|V_w \cap \partial\tau| > 2$.

Finally, we prove (3). We start by showing that for $w \in \mathcal{W}_{k/2+d}^k$ the graph G_w is a tree. Indeed, the graph G_w is connected by definition and it cannot contain a loop since this would imply that one of the d -cells does not add a new 0-cell in its first appearance. Since G_w is a tree that comes with an additional path covering its vertices and edges we can think of G_w as a rooted planar tree by declaring

the first letter of w to be its root and choosing a planar embedding of it that will make the path generated by w into a clockwise exploration path of the tree (see [AGZ10, Lemma 2.1.6]).

Finally, we turn to evaluate $|\mathcal{W}_{k/2+d}^k|$. It is well known, see for example [AGZ10, Lemma 2.1.6], that the set rooted planar trees with $\frac{k}{2}$ edges is in bijection with Dyck paths of length k and is thus of size $\mathcal{C}_{k/2}$. We will show that the set $\mathcal{W}_{k/2+d}^k$ is in bijection with the set of rooted, labeled planar trees with $\frac{k}{2}$ edges or more formally the set of rooted planar trees with a label from $\llbracket d \rrbracket$ attached to each of the tree edges. Since the total number of possible labelings is $d^{k/2}$ it follows that $|\mathcal{W}_{k/2+d}^k| = \mathcal{C}_{k/2} d^{k/2}$, thus completing the proof.

It is hence left to construct the aforementioned bijection. To this end, we label the 0-cells of X by the numbers in $\llbracket n \rrbracket$ and associate with every j -cell $\sigma = \{\sigma^0, \dots, \sigma^j\} \in \binom{\llbracket n \rrbracket}{j+1}$ an ordering $\langle \sigma \rangle = (\sigma^{i_0}, \dots, \sigma^{i_j})$ such that $\sigma^{i_0} < \sigma^{i_1} < \dots < \sigma^{i_j}$. In order to make the bijection simpler to write we choose to work with the specific choice of orientation X_+^j associated with the above ordering defined by $X_+^j = \{[\sigma^0, \dots, \sigma^j] : \sigma^0 < \dots < \sigma^j\}$. Finally, we fix a special representative $w \in \mathcal{W}_{k/2+d}^k$ by requiring that the 0-cells in w are $\{1, \dots, \frac{k}{2} + d\}$ and that they appear along the path w in increasing order.

Given $w \in \mathcal{W}_{k/2+d}^k$ define its rooted, labeled planar tree (G_w, ℓ_w) by letting G_w be the graph of the word w as defined in Definition 3.6 with the planar embedding that makes w into a clockwise exploration path. The labeling $\ell_w : E_w \rightarrow \llbracket d \rrbracket$ is then defined as follows: For an edge $\{\sigma, \sigma'\} \in E_w$ let $1 \leq i \leq k$ be the first integer such that $\{\sigma_i, \sigma_{i+1}\} = \{\sigma, \sigma'\}$. Then we define $\ell(\{\sigma, \sigma'\}) = j$ if and only if the $(d-2)$ -cell $\sigma_i \cap \sigma_{i+1}$ is obtained from σ_i by deleting the j -th smallest 0-cell in it (see Figure 3.3 for an illustration).

In the other direction, given a rooted, labeled planar tree $(G = (V, E), \ell : E \rightarrow \llbracket d \rrbracket)$, define $w = \sigma_1 \dots \sigma_k \sigma_1 \in \mathcal{W}_{k/2+d}^k$ by following procedure:

- Associate with the root of the tree the letter $\sigma_1 = [1, \dots, d]$.
- Following the exploration path of the tree clockwise, if in the i -th step of the exploration the explored vertex appeared before, say in the j -th step, define $\sigma_i = \sigma_j$. If the i -th explored vertex is a new vertex then define σ_i to be $\sigma_i = \sigma_{i-1} \setminus \sigma_{i-1}^{\ell(\{\sigma_{i-1}, \sigma_i\})} \cup \{m\}$, where $m = |\sigma_1 \cup \sigma_2 \cup \dots \cup \sigma_{i-1}| + 1$. That is, σ_i is obtained from σ_{i-1} by deleting the $\ell(\{\sigma_{i-1}, \sigma_i\})$ -th smallest number in σ_{i-1} (with $\ell(\{\sigma_{i-1}, \sigma_i\})$ being the label of the currently explored edge) and adding a new 0-cell with the smallest number that did not appear so far.

One can easily verify that these two maps are inverses of each other, thus completing the proof. \square

4 Proof of Theorem 2.5

Since the eigenvalues of $(nq)^{-1/2}(A + pdI)$ and $(nq)^{-1/2}A$ are simply shifted by $\frac{dp}{\sqrt{nq}}$, which by assumption tends to zero as $n \rightarrow \infty$, it is enough to prove the semicircle law for the matrix $(nq)^{-1/2}(A + pdI)$.

Using (3.1) and (3.4) we can write

$$\frac{1}{\sqrt{nq}}(A + pdI) = H + \frac{p}{\sqrt{nq}}(\mathbb{A} + dI). \quad (4.1)$$

Lemma 4.1 ([GW14] Lemma 8). *The eigenvalues of \mathbb{A} are $n - d$ with multiplicity $\binom{n-1}{d-1}$ and $-d$ with multiplicity $\binom{n-1}{d}$.*

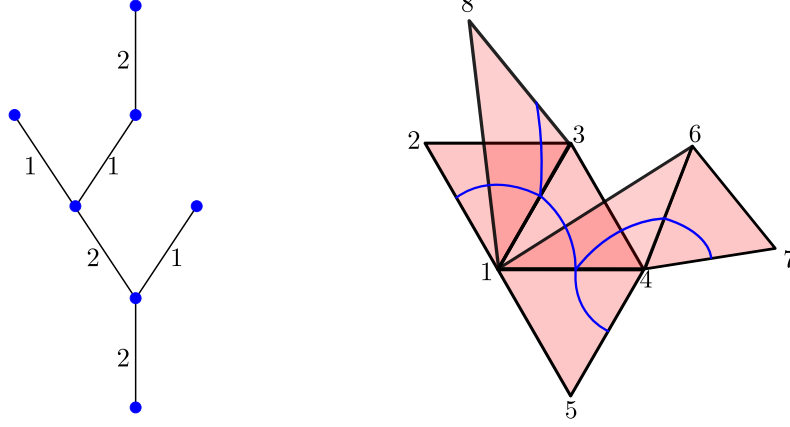


Figure 3.3: (a) An illustration of the labeled tree associated with the word $w = [1, 2][1, 3][1, 4][4, 5][1, 4][4, 6][4, 7][4, 6][1, 4][1, 3][3, 8][1, 3][1, 2]$. (b) The path of the word w on the complex.

It follows that the matrix $\mathbb{A} + dI$ has rank $\binom{n-1}{d-1}$ and therefore by Weyl's interlacing inequalities we have

$$\lambda_i(H) \leq \lambda_i \left(\frac{1}{\sqrt{nq}}(A + pdI) \right) \leq \lambda_{i+\binom{n-1}{d-1}}(H), \quad \forall i \in \mathbb{Z}, \quad (4.2)$$

where for a self-adjoint, $N \times N$ matrix B we define $\lambda_i(B) = -\infty$ for $i < 1$ and $\lambda_i(B) = +\infty$ for $i > N$.

For an $N \times N$ matrix B denote by $\kappa_B(x) = L_B((-\infty, x])$ the cumulative distribution function of its eigenvalues. Let f be a smooth bounded function with a bounded derivative. Then, using integration by parts we get

$$\begin{aligned} \int_{\mathbb{R}} f(E) d\kappa_{(nq)^{-1/2}(A+pdI)}(E) - \int_{\mathbb{R}} f(E) d\kappa_H(E) \\ = \int f'(E) \left(\kappa_H(E) - \kappa_{(nq)^{-1/2}(A+pdI)}(E) \right) dE. \end{aligned} \quad (4.3)$$

Due to (4.2) the right hand side is bounded by

$$\frac{\|f'\|_{\infty}}{N} \binom{n-1}{d-1} = O \left(\frac{\|f'\|_{\infty}}{n} \right),$$

and thus (4.3) goes to zero as $n \rightarrow \infty$.

It now follows by an approximation argument that

$$\lim_{n \rightarrow \infty} \left| \int_{\mathbb{R}} f(E) d\kappa_{(nq)^{-1/2}(A+pdI)}(E) - \int_{\mathbb{R}} f(E) d\kappa_H(E) \right| = 0 \quad (4.4)$$

for every bounded continuous function f . Therefore, by Theorem 3.1

$$\lim_{n \rightarrow \infty} L_{(nq)^{-1/2}(A+pdI)} = \lim_{n \rightarrow \infty} L_H = \nu_d$$

almost surely (where the limits are in distribution), and Theorem 2.5 follows. \square

5 Bounding the norm of H

In this section we prove the following bound on the norm of the matrix H defined in (3.1).

Theorem 5.1. *Suppose that $nq \geq 2$. Then for every $\xi > 0$ we have*

$$\mathbb{P}(\|H\| > 2\sqrt{d} + \xi) \leq \mathcal{E}(\xi),$$

where we introduced the error probability

$$\mathcal{E}(\xi) \equiv \mathcal{E}_{n,p,d}(\xi) := \frac{2 \left(1 + \frac{\xi}{2\sqrt{d}}\right)^2}{(d-1)!} \exp \left(d \log n - \left(\frac{2}{3} \log \left(1 + \frac{\xi}{2\sqrt{d}} \right) \right)^{3/2} \left(\frac{nq}{d} \right)^{1/4} \right). \quad (5.1)$$

Remark 5.2. In particular, if $\lim_{n \rightarrow \infty} \frac{nq}{\log^4 n} = \infty$ it follows that from the Borel-Cantelli lemma that for every $\varepsilon > 0$ we have almost surely $\limsup_{n \rightarrow \infty} \|H\| < 2\sqrt{d} + \varepsilon$. Since ε was arbitrary we conclude that $\limsup_{n \rightarrow \infty} \|H\| \leq 2\sqrt{d}$ almost surely. Observing that the semicircle law (Theorem 3.1) implies that almost surely $\liminf_{n \rightarrow \infty} \|H\| \geq 2\sqrt{d}$ we obtain that almost surely $\lim_{n \rightarrow \infty} \|H\| = 2\sqrt{d}$.

The proof of Theorem 5.1 uses a Füredi-Komlós-type argument; see [FK81, AGZ10] for a presentation of the classical Füredi-Komlós argument for matrices with independent entries. The main work is to estimate the number of equivalence classes in \mathcal{W}_s^k . This estimate is given in the following result.

Proposition 5.3. *For every $k \geq 0$ and $d \leq s \leq \lfloor \frac{k}{2} \rfloor + d$*

$$|\mathcal{W}_s^k| \leq d(2\sqrt{d})^k \sum_{m=0}^{k-2(s-d)} \frac{\left(\frac{\sqrt{d}}{2} k^3\right)^m}{m!}.$$

Equipped with Proposition 5.3, whose proof is postponed, the proof of Theorem 5.1 is standard.

Proof of Theorem 5.1. Using (3.10) we can write for even k

$$\begin{aligned} \mathbb{E} \left[\int_{\mathbb{R}} x^k L_H(dx) \right] &= \sum_{s=d}^{\frac{k}{2}+d} \frac{B_{n,s,d}}{N(nq)^{k/2}} \sum_{w \in \mathcal{W}_s^k} \prod_{\tau \in X^d} \mathbb{E} \left[(\chi - p)^{N_w(\tau)} \right] \text{sgn}(w, \tau) \\ &\stackrel{(1)}{\leq} \sum_{s=d}^{\frac{k}{2}+d} \frac{n^{s-d} |\mathcal{W}_s^k|}{(nq)^{k/2}} \sup_{w \in \mathcal{W}_s^k} \left| \prod_{\tau \in X^d} \mathbb{E} \left[(\chi - p)^{N_w(\tau)} \right] \right| \\ &\stackrel{(2)}{\leq} \sum_{s=d}^{\frac{k}{2}+d} \frac{|\mathcal{W}_s^k|}{(nq)^{k/2+d-s}} \\ &\stackrel{(3)}{\leq} d(2\sqrt{d})^k \sum_{m=0}^k \frac{\left(\frac{\sqrt{d}}{2} k^3\right)^m}{(nq)^{k/2} m!} \frac{(nq)^{\frac{k-m}{2}+1} - 1}{nq - 1} \\ &\leq d(2\sqrt{d})^k \frac{nq}{nq - 1} \sum_{m=0}^k \frac{1}{m!} \left(\sqrt{\frac{d}{4nq}} k^3 \right)^m \\ &\stackrel{(4)}{\leq} 2d(2\sqrt{d})^k \exp \left(\sqrt{\frac{d}{4nq}} k^3 \right), \end{aligned}$$

where for (1) we used the fact that $N^{-1}B_{n,s,d} \leq n^{s-d}$, for (2) we used the fact that the supremum is bounded by q^{s-d} (see (3.12)), for (3) we used Proposition 5.3, and for (4) we used the assumption $nq \geq 2$.

Therefore, it follows from the Markov's inequality that for every $\xi > 0$ and k even

$$\begin{aligned} & \mathbb{P}(\|H\| > 2\sqrt{d} + \xi) \\ & \leq \mathbb{P}\left(\sum_{i=1}^N (\lambda_i(H))^k > (2\sqrt{d} + \xi)^k\right) = \mathbb{P}\left(\int_{\mathbb{R}} x^k L_H(dx) > \frac{(2\sqrt{d} + \xi)^k}{N}\right) \\ & \leq \frac{N \mathbb{E}[\int_{\mathbb{R}} x^k L_H(dx)]}{(2\sqrt{d} + \xi)^k} \leq \frac{2n^d}{(d-1)!} \left(\frac{2\sqrt{d}}{2\sqrt{d} + \xi}\right)^k \exp\left(\sqrt{\frac{d}{4nq}} k^3\right) \\ & \leq \frac{2}{(d-1)!} \exp\left(d \log n + \left[\sqrt{\frac{d}{4nq}} k^2 - \log\left(1 + \frac{\xi}{2\sqrt{d}}\right)\right] k\right). \end{aligned}$$

The claim then follows by choosing $k = k(n)$ to be the largest even integer that is smaller than $\left(\frac{2}{3\sqrt{d}} \log\left(1 + \frac{\xi}{2\sqrt{d}}\right)\right)^{1/2} (nq)^{1/4}$. \square

The rest of this section is devoted to the proof of Proposition 5.3. We start with the following definitions, which provide the correct generalization of the classical definition of FK words [AGZ10] to the case $d > 1$.

Definition 5.4 (Sentences). An (n, d) -sentence (or shortly a *sentence*) is a finite sequence of words (w_1, w_2, \dots, w_m) at least one word long. The *length of a sentence* a is the sum of the lengths of its words. Note that unlike for words, in general the length is not the same as the number of crossings minus one.

Definition 5.5 (Support of a sentence). For a sentence $a = (w_1, \dots, w_m)$ we denote by $\text{supp}_0(a)$ and $\text{supp}_d(a)$ the union of $\text{supp}_0(w_i)$ respectively $\text{supp}_d(w_i)$ over $1 \leq i \leq m$.

Definition 5.6 (Graph of a sentence). Given a sentence $a = (w_1, \dots, w_m)$ with $w_i = \sigma_{i,1}\sigma_{i,2} \dots \sigma_{i,\ell_i}$ we define the graph associated with it $G_a = (V_a, E_a)$ by

$$V_a = \{\sigma_{i,j} : 1 \leq i \leq m, 1 \leq j \leq \ell_i\}$$

and

$$E_a = \{\{\sigma_{i,j}, \sigma_{i,j+1}\} : 1 \leq i \leq m, 1 \leq j \leq \ell_i - 1\}.$$

As in the word case the sentence a induces a sequence of paths on the graph G_a which together cover all of its vertices and edges. Also, the paths induces an ordering on the $(d-1)$ -cells, 0-cells and d -cells associated with the graph by following the words according to their order in the sentence and following the usual order inside each word. As in the word case we define $N_a(e) = \sum_{i=1}^m N_{w_i}(e)$ to be the number of times the edge $e \in E_a$ is crossed in the sentence a , let $E_a(\tau) = \{\{\sigma, \sigma'\} \in E_a : \sigma \cup \sigma' = \tau\}$ be the set of edges associated with the d -cell τ and define $N_a(\tau) = \sum_{e \in E_a(\tau)} N_a(e)$ to be the total number of crossings of τ .

Definition 5.7 (Wigner words). A closed word w of length $k+1 \geq 1$ is called a *Wigner word* if either $k = 0$ or $k > 0$ is even and w is equivalent to an element of $\mathcal{W}_{k/2+d}^k$.

Definition 5.8 (FK words). A word w is called an *FK word* if the graph G_w associated with w is a tree, $N_w(e) \leq 2$ for every $e \in E_w$, and for each $\tau \in X^d$ there is at most one edge $e \in E_w(\tau)$ such that $N_w(e) = 2$. More generally, a sentence $a = (w_1, \dots, w_m)$ is called an *FK sentence* if the graph G_a is a tree, $N_a(e) \leq 2$ for every $e \in E_a$, for each $\tau \in X^d$ there is at most one edge $e \in E_a(\tau)$ such that $N_a(e) = 2$, and for every $1 \leq i \leq m-1$ the first letter of w_{i+1} belongs to one of the words w_1, \dots, w_i . Two FK sentences $a = (w_1, \dots, w_m)$ and $a' = (w'_1, \dots, w'_{m'})$ are called *equivalent* if $m = m'$ and there exists a permutation π on $X^0 = V$ such that $\pi(w_i) = \pi(w'_i)$ for $1 \leq i \leq m$ (see Definition 3.4 for the meaning of $\pi(w_i)$).

Definition 5.9 (Word parsing). A *parsing* of a word $w = \sigma_1 \dots \sigma_k$ is a sentence $a = (w_1, \dots, w_m)$ such that the concatenation of its words yields w . We say that $w = \sigma_1 \dots \sigma_k$ is parsed at time i (or parsed in $\sigma_i \sigma_{i+1}$) in the parsing a if σ_i and σ_{i+1} do not belong to the same word in a .

The key definition above is that of the FK word, which is, as it turns out, the correct generalization of the definition in [AGZ10, Section 2.1.6] to $d > 1$.

Before starting the formal proof of the lemma let us explain its general scheme; see also [AGZ10] for more explanations on the structure of the Füredi-Komlós argument for matrices with independent entries. The proof starts by showing that each word $w \in \mathcal{W}_s^k$ can be parsed into an FK sentence. Since the original word can be read from its parsing simply by concatenating the words, it follows that it is enough to bound the number of such FK sentences. This is obtained via the following three steps. (a) Bounding the number of words in the FK sentence. (b) Bounding the number of ways to choose the equivalence classes of the FK words for the FK sentence. (c) Bounding the number of ways one can “glue” the chosen FK words together in order to obtain an FK sentence. The first bound, (a), which is obtained for $d = 1$ by a simple observation (see [AGZ10]), requires a new argument for $d > 1$, which is similar in spirit to the one used to prove Lemma 3.11(2). The second bound, (b), is obtained by showing that each FK word can be parsed to a sentence which is comprised of disjoint Wigner words and then applying Lemma 3.11(3) to bound their number. Finally, the bound (c) on the number of ways to “glue” the FK words together is obtained by showing that a “good gluing” implies the existence of joint geodesic (see the proof of Lemma 5.12).

Proof of Proposition 5.3 assuming three lemmas. Given a closed word w of length $k + 1$ we define its FK parsing a_w as follows: Declare an edge e of the associated graph G_w *new* if for some index $1 \leq i \leq k$ we have $e = \{\sigma_i, \sigma_{i+1}\}$ and $\sigma_{i+1} \notin \{\sigma_1, \dots, \sigma_i\}$. If an edge e is not *new*, then it is *old* (note that old edges are the ones that create a loop in G_w in their first crossing). Define a_w to be the parsing obtained from w by parsing in $\sigma_i \sigma_{i+1}$ if one of the following occurs:

- $\{\sigma_i, \sigma_{i+1}\}$ is an old edge of G_w .
- $\{\sigma_i, \sigma_{i+1}\}$ is a third or subsequent crossing of the edge $\{\sigma_i, \sigma_{i+1}\}$.
- $\{\sigma_i, \sigma_{i+1}\}$ is a second crossing of the edge $\{\sigma_i, \sigma_{i+1}\}$ and there exists an edge $\{\sigma, \sigma'\} \in E_w(\sigma_i \cup \sigma_{i+1})$ which is crossed twice in $\sigma_1 \sigma_2 \dots \sigma_i$.

Since this parsing eliminates all loops in G_w , third and subsequent visits to edges and leaves in each d -cell at most one edge which is crossed twice the resulting sentence a_w is indeed an FK sentence.

As indicated before, every word w can be recovered from its FK sentence a_w . Therefore it is enough to bound the number of such FK sentences. The following three lemmas contain the required bounds for this purpose.

Lemma 5.10. *For every $w \in \mathcal{W}_s^k$ the number of words m_w in the FK sentence a_w satisfies*

$$1 \leq m_w \leq k + 1 - 2(s - d).$$

Lemma 5.11. *There are at most $\frac{\sqrt{d}}{2}(2\sqrt{d})^k$ equivalence classes of FK words of length k .*

Lemma 5.12. *Let b be an FK sentence comprised of $m - 1$ words of total length l and z an FK word. Then there are at most l^2 FK words w equivalent to z such that (b, w) is an FK sentence. (Informally, there are at most l^2 ways to glue z to b so that the result is an FK sentence.)*

Assuming the last three lemmas, we turn to complete the proof. Assume first that $m_w = m$. There are $\binom{k}{m-1}$ ways to choose m -tuples (l_1, \dots, l_m) of positive integers summing to $k + 1$ which are the length of the words in the sentence a_w . Given the lengths of the words, by Lemma 5.11 there are at most $\prod_{i=1}^m \left(\frac{\sqrt{d}}{2}(2\sqrt{d})^{l_i}\right) = \left(\frac{\sqrt{d}}{2}\right)^m (2\sqrt{d})^{k+1}$ ways to choose equivalence classes for each of these FK words. Finally, due to Lemma 5.12, there are at most $l_1^2 (l_1 + l_2)^2 \cdots (l_1 + \dots + l_{m-1})^2 \leq k^{2(m-1)}$ ways to choose representatives in order to obtain an FK sentence.

Combining all of the above, we find that there are at most

$$\binom{k}{m-1} \left(\frac{\sqrt{d}}{2}\right)^m (2\sqrt{d})^{k+1} k^{2(m-1)}$$

equivalence classes of FK sentences with m words.

Using now Lemma 5.10 we conclude that the number of equivalence classes of FK sentences of length $k + 1$ is bounded by

$$\begin{aligned} \sum_{m=1}^{k+1-2(s-d)} \binom{k}{m-1} \left(\frac{\sqrt{d}}{2}\right)^m (2\sqrt{d})^{k+1} k^{2(m-1)} \\ \leq \sum_{m=1}^{k+1-2(s-d)} \left(\frac{\sqrt{d}}{2}\right)^m (2\sqrt{d})^{k+1} \frac{k^{3(m-1)}}{(m-1)!} = d(2\sqrt{d})^k \sum_{m=0}^{k-2(s-d)} \frac{\left(\frac{\sqrt{d}}{2}k^3\right)^m}{m!}, \end{aligned}$$

thus completing the proof. \square

Next we turn the proof of Lemmas 5.10–5.12.

Proof of Lemma 5.10. Fix $w \in \mathcal{W}_s^k$ and let a_w be the parsing of w into an FK sentence consisting of m_w words. Denoting by r_w the number of crossings along a_w , we have $k + 1 = m_w + r_w$. Thus, it is enough to show that $r_w \geq 2(s - d)$. Since $|\text{supp}_0(w)| = s$ it follows that there are $(s - d)$ crossings in w , denoted by $(\sigma_{i(j)}\sigma_{i(j)+1})_{j=1}^{s-d}$, such that $\sigma_{i(j)+1}$ contains a 0-cell that does not belong to $\sigma_1 \cup \dots \cup \sigma_{i(j)}$. Without loss of generality, we assume that the 0-cells are indexed by $\llbracket n \rrbracket$ and that they are discovered in increasing order. Therefore in the crossing $\sigma_{i(j)}\sigma_{i(j)+1}$ we discover the 0-cell $d + i(j)$. Since $\sigma_{i(j)+1}$ contains a new 0-cell that did not appear in $\sigma_1 \cup \dots \cup \sigma_{i(j)}$ we must have that $\sigma_{i(j)}\sigma_{i(j)+1}$ is a first crossing of the edge $\{\sigma_{i(j)}, \sigma_{i(j)+1}\}$ and that it does not create a loop (i.e., it is a new edge). Consequently, $\sigma_{i(j)}\sigma_{i(j)+1}$ for $1 \leq j \leq s - d$ are all crossings in a_w as well. In addition, the crossings $\sigma_{i(j)}\sigma_{i(j)+1}$ for $1 \leq j \leq s - d$ are all disjoint since the d -cells associated with them $\sigma_{i(j)} \cup \sigma_{i(j)+1}$ are disjoint. Thus, $r_w \geq s - d$.

Next, for $1 \leq j \leq s - d$ denote by $\sigma_{k(j)}\sigma_{k(j)+1}$ the second crossing in w of the d -cell $\sigma_{i(j)} \cup \sigma_{i(j)+1}$ in w . Each of the crossings $\sigma_{k(j)}\sigma_{k(j)+1}$ which is also a crossing in a_w increases r_w by one and so we only need to find an additional crossing of a_w to replace those crossings $\sigma_{k(j)}\sigma_{k(j)+1}$ in w which are not crossings in a_w . Fix $1 \leq j \leq s - d$ such that $\sigma_{k(j)}\sigma_{k(j)+1}$ is not a crossing in a_w . Since $\sigma_{k(j)}\sigma_{k(j)+1}$ is the second crossing of the d -cell $\sigma_{i(j)} \cup \sigma_{i(j)+1}$ this can only happen if the edge $\{\sigma_{k(j)}, \sigma_{k(j)+1}\}$ is an old edge which in particular implies that $\{\sigma_{k(j)}, \sigma_{k(j)+1}\} \neq \{\sigma_{i(j)}, \sigma_{i(j)+1}\}$, i.e., there are at least 3 distinct $(d - 1)$ -cells in the boundary of the d -cell $\sigma_{i(j)} \cup \sigma_{i(j)+1}$ appearing along w . Let $\tilde{\sigma}_j$ denote the first $(d - 1)$ -cell in the boundary of $\sigma_{i(j)} \cup \sigma_{i(j)+1}$ that appears in w and is not $\sigma_{i(j)}$ or $\sigma_{i(j)+1}$. Finally,

let $\sigma_{l(j)}\sigma_{l(j)+1}$ denote the first crossing in w such that $\sigma_{l(j)+1} = \tilde{\sigma}_j$. Since $\sigma_{l(j)+1}$ is the first appearance of $\tilde{\sigma}_j$, the crossing $\sigma_{l(j)}\sigma_{l(j)+1}$ is not of an old edge and is a first visit to the edge $\{\sigma_{l(j)}, \sigma_{l(j)+1}\}$. Consequently, it is also a crossing in a_w .

For $1 \leq j \leq s - d$ define

$$r(j) = \begin{cases} k(j) & \text{if } \sigma_{k(j)}\sigma_{k(j)+1} \text{ is a crossing in } a_w \\ l(j) & \text{if } \sigma_{k(j)}\sigma_{k(j)+1} \text{ is not a crossing in } a_w \end{cases}.$$

In order to show that $r_w \geq 2(s - d)$ and thus to complete the proof it is thus left to show that the crossing times $\{i(j)\}_{j=1}^{s-d} \cup \{r(j)\}_{j=1}^{s-d}$ are distinct. This is indeed the case, as can be seen by the following observations:

- It was already observed before that the times $\{i(j)\}_{j=1}^{s-d}$ are all distinct as the associated crossings $\sigma_{i(j)}\sigma_{i(j)+1}$ form distinct d -cells. Similarly, for every $1 \leq j_1, j_2 \leq s - d$ the crossing times $i(j_1)$ and $r(j_2)$ are distinct. Indeed, $i(j_1)$ is the first crossing time in which the 0-cell $j_1 + d$ is observed. Since $r(j_2)$ cannot be a crossing time in which a new 0-cell is observed, it follows that the times $i(j_1)$ and $r(j_2)$ must be distinct.
- Finally, we claim that the times $\{r(j)\}_{j=1}^{s-d}$ are distinct. Indeed, assume that for some $1 \leq j_1 < j_2 \leq s - d$ we have $r(j_1) = r(j_2)$.
 - If $r(j_1) = k(j_1)$ and $r(j_2) = k(j_2)$, then $r(j_1) \neq r(j_2)$ since the d -cells $\sigma_{k(j_1)} \cup \sigma_{k(j_1)+1} = \sigma_{i(j_1)} \cup \sigma_{i(j_1)+1}$ and $\sigma_{k(j_2)} \cup \sigma_{k(j_2)+1} = \sigma_{i(j_2)} \cup \sigma_{i(j_2)+1}$ associated with the crossings are distinct and thus in particular so are the crossing times.
 - If $r(j_1) = k(j_1)$ and $r(j_2) = l(j_2)$, then $r(j_1) \neq r(j_2)$ since the d -cell $\sigma_{l(j_2)} \cup \sigma_{l(j_2)+1}$ contains the 0-cell $d + j_2$ which by definition does not belong to the d -cell $\sigma_{k(j_1)} \cup \sigma_{k(j_1)+1} = \sigma_{i(j_1)} \cup \sigma_{i(j_1)+1}$.
 - If $r(j_1) = l(j_1)$ and $r(j_2) = k(j_2)$, then $r(j_1) \neq r(j_2)$ (see Figure 5.1). Indeed, assume that $r(j_1) = r(j_2)$. Since the d -cell $\sigma_{l(j_1)} \cup \sigma_{l(j_1)+1}$ is the d -cell in which the 0-cell $d + j_2$ appears for the first time we must have that $d + j_2$ belongs to $\sigma_{l(j_1)} \cup \sigma_{l(j_1)+1}$. In addition, since $\sigma_{l(j_1)+1}$ belongs to $\sigma_{i(j_1)} \cup \sigma_{i(j_1)+1}$ and $j_1 < j_2$, it follows that the 0-cell $d + j_2$ does not belong to $\sigma_{l(j_1)+1}$ and so it must be in $\sigma_{l(j_1)}$. However, this implies that $\sigma_{k(j_2)} \cup \sigma_{k(j_2)+1} = \sigma_{i(j_2)} \cup \sigma_{i(j_2)+1}$ is not the first d -cell containing the 0-cell $d + j_2$, which contradicts its definition.
 - If $r(j_1) = l(j_1)$ and $r(j_2) = l(j_2)$, then $r(j_1) \neq r(j_2)$. Indeed, we must have that $\sigma_{l(j_1)+1}$, which belongs to the boundary of $\sigma_{i(j_1)} \cup \sigma_{i(j_1)+1}$, only contains 0-cells that appeared until the crossing of $\sigma_{i(j_1)} \cup \sigma_{i(j_1)+1}$, i.e., only 0-cells from $1, 2, \dots, j_1 + d$. However, $\sigma_{l(j_2)+1}$ must contain the 0-cell $d + j_2$ by definition. This leads to a contradiction since $j_1 < j_2$.

This concludes the proof. \square

Proof of Lemma 5.11. We start by showing that each FK word can be written in a unique way as a concatenation of disjoint Wigner words. Let $w = \sigma_1 \dots \sigma_l$ be an FK word of length l . Let $\{\sigma_{i(j)}, \sigma_{i(j)+1}\}_{j=1}^r$ be the edges in the graph G_w which are crossed precisely once by the path w . Defining $i(0) = 0$ and $w_j = \sigma_{i(j-1)+1} \dots \sigma_{i(j)}$, we now claim that the words w_j are closed and disjoint and that G_{w_j} is a tree in which every d -cell, and every edge are crossed precisely twice, i.e., the words w_j are disjoint Wigner words. Indeed, denote the above parsing of w by b_w . Since G_w is a tree, the graph G_{b_w} is a forest and the paths generated by b_w on it cross each edge precisely twice. In addition, since in an FK word there is at most one edge inside each d -cell which is crossed twice, it follows that each d -cell τ which is

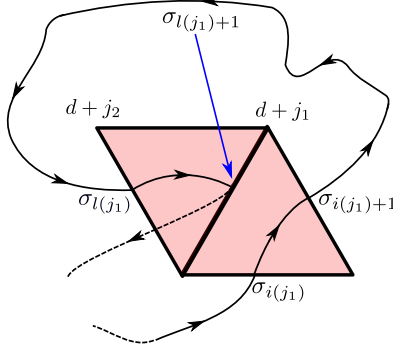


Figure 5.1: Illustration for the case $r(j_1) = l(j_1)$ and $r(j_2) = k(j_2)$. Note that the in order for $r(j_1)$ to be the same as $r(j_2)$ the path must reach the $(d-1)$ -cell $\sigma_{l(j_2)+1}$ (which does not contain the 0-cell $d + j_2$) through a d -cell that contains the 0-cell $d + j_2$. However, this implies that this d -cell is not the first d -cell in which the 0-cell $d + j_2$ appears.

crossed in b_w is crossed precisely twice; both times on the unique edge $e \in E_w(\tau)$ such that $N_w(e) = 2$. Finally, note that the words w_j are closed since each edge in their graph is crossed precisely twice and it contains no loops as a subgraph of G_{b_w} . Thus each word must be associated with a unique connected component of G_{b_w} and its path must cross each of the edges and d -cells of this component twice.

From this point on the proof proceeds in the same way as in the case $d = 1$ (see [AGZ10, Lemma 2.1.24]), and we include the conclusion of the proof, adapted from [AGZ10, Lemma 2.1.24], for the reader's convenience. The aforementioned decomposition is unique and one concludes that with N_k denoting the number of equivalence classes of FK words of length k we get

$$\begin{aligned} \sum_{k=1}^{\infty} N_k z^k &= \sum_{r=1}^{\infty} \sum_{l_1, l_2, \dots, l_r \in 2\mathbb{N}} \prod_{j=1}^r z^{l_j+1} |\mathcal{W}_{l_j/2+d}^{l_j}| = \sum_{r=1}^{\infty} \left(z + \sum_{l=1}^{\infty} z^{2l+1} |\mathcal{W}_{l+d}^{2l}| \right)^r \\ &= \sum_{r=1}^{\infty} \left(z + \sum_{l=1}^{\infty} z^{2l+1} C_l d^l \right)^r, \end{aligned}$$

where for the last step we used Lemma 3.11. Consequently, using the generating function of the Catalan numbers, for $|z| < \frac{1}{2\sqrt{d}}$ we have

$$\sum_{k=1}^{\infty} N_k z^k = \sum_{r=1}^{\infty} \left(\frac{1 - \sqrt{1 - 4dz^2}}{2dz} \right)^r = \frac{1}{2} \left(-1 + \frac{1 + 2dz}{\sqrt{1 - 4dz^2}} \right).$$

Since $\frac{1}{\sqrt{1-t}} = \sum_{k=0}^{\infty} \frac{1}{4^k} \binom{2k}{k} t^k$ for $|t| < 1$ it follows that

$$\sum_{k=1}^{\infty} N_k z^k = -\frac{1}{2} + \frac{1}{2}(1 + 2dz) \sum_{k=1}^{\infty} d^k \binom{2k}{k} z^{2k}$$

and thus

$$|N_k| \leq \frac{\sqrt{d}}{2} (2\sqrt{d})^k. \quad \square$$

With our definition of FK sentences, Lemma 5.12 follows by the same argument as in the case $d = 1$ (see [AGZ10, Lemma 2.1.25]). For the reader's convenience we include the appropriately adapted version of this argument here. We start with some additional definitions. Recall from the proof of

Lemma 5.11 that any FK word w can be written in a unique way as a concatenation of disjoint Wigner words $b_w = (w_1, \dots, w_r)$. With $\sigma_{i,1}$ denoting the first (and last) letter of w_i , define the *skeleton* of the FK word w to be the word $\tilde{w} = \sigma_{1,1}\sigma_{2,1} \dots \sigma_{r,1}$. Finally, for an FK sentence a with a graph G_a define a new graph $\tilde{G}_a = (\tilde{V}_a, \tilde{E}_a)$ by setting $\tilde{V}_a = V_a$ and $\tilde{E}_a = \{e \in E_a : N_a(e) = 1\}$. Since G_a is a tree and the edges removed from it are the one corresponding to edges of the Wigner words G_{w_j} , the resulting graph \tilde{G}_a is a forest.

Proof of Lemma 5.12. As averted above, the proof is almost identical to that of [AGZ10, Lemma 2.1.25]. We begin with the following observation. Suppose b is an FK sentence with $m - 1$ words and that w is an FK word equivalent to z such that (b, w) is an FK sentence. Denote the skeleton of w by $\tilde{w} = \sigma_{j_1}\sigma_{j_2} \dots \sigma_{j_l}$, so that σ_{j_1} is a letter in b by definition. Let l be the largest index such that σ_{j_l} is a letter in b and set $w' = \sigma_{j_1} \dots \sigma_{j_l}$. Then $V_b \cap V_w = V_{w'}$ and w' is a geodesic in \tilde{G}_b .

Assuming the observation holds we finish the proof. The number of ways to choose an FK word w equivalent to z such that (b, w) is an FK sentence is bounded by the number of ways to choose a geodesic in its forest \tilde{G}_b . The number of ways to choose a geodesic inside a forest is bounded by the number of ways to choose its endpoints (since the geodesic is unique) which in turn is bounded by l^2 .

The proof of the observation follows from the same argument that is used in the graph case $d = 1$; see [AGZ10, Lemma 2.1.25]. Suppose a is an FK sentence. Then G_a is a tree, and since the Wigner words composing w are disjoint, w' is the unique geodesic in $G_w \subset G_a$ connecting σ_{j_1} to σ_{j_l} . But w' visits only edges of G_b that have been visited exactly once by the words constituting b , for otherwise (b, w) would not be an FK sentence (that is, a *comme* would need to be inserted to split w). Thus $E_{w'} \subset \tilde{E}_b$. Since w is an FK word, $\tilde{E}_w = E_{\tilde{w}}$. Since a is an FK sentence $E_b \cap E_w = \tilde{E}_b \cap \tilde{E}_w$. Thus $E_b \cap E_w = E_{w'}$. But, now recall that $G_a, G_b, G_w, G_{w'}$ are all trees and hence

$$\begin{aligned} |V_a| &= 1 + |E_a| = 1 + |E_b| + |E_w| - |E_b \cap E_w| = 1 + |E_b| + |E_w| - |E_{w'}| \\ &= 1 + |E_b| + 1 + |E_w| - 1 - |E_{w'}| = |V_b| + |V_w| - |V_{w'}|. \end{aligned}$$

Since $|V_b| + |V_w| - |V_b \cap V_w| = |V_a|$, it follows that $|V_{w'}| = |V_b \cap V_w|$. Since $V_{w'} \subset V_b \cap V_w$ one concludes that $V_{w'} = V_b \cap V_w$, as claimed. \square

6 Proof of Theorem 2.1

We start this section with the following stronger version of Theorem 2.1.

Theorem 6.1. *Assume $d \geq 2$ and $nq \geq 10^3 d^2$.*

- (1) *For every $\xi > 0$, the $\binom{n-1}{d}$ smallest eigenvalues of the matrix A are within the interval $-pd + \sqrt{nq}[-2\sqrt{d} - \xi, 2\sqrt{d} + \xi]$ with probability at least $1 - \mathcal{E}(\xi)$ (see (5.1)).*
- (2) *For every $\xi > 0$, if $nq \geq d(2d + 2\xi)^6 n^{-1} \log^6 n$, then the remaining $\binom{n-1}{d-1}$ eigenvalues of A are inside the interval $np + [-\Gamma(\xi, n), \Gamma(\xi, n)]$ with probability at least $1 - \mathcal{E}((\sqrt{6} - 2)\sqrt{d}) - \mathcal{E}(\xi)$ where*

$$\mathcal{E}(\xi) \equiv \mathcal{E}_{n,d}(\xi) := \frac{4e^3 d^{5/2}}{(d-1)!} \exp(5 \log(2d + 2\xi) + 5 \log \log n - \xi \log n)$$

and

$$\Gamma(\xi, n) := 6d + \frac{200d^{3/2}}{\sqrt{nq} - 24d} + 100d^{7/2}(d + \xi)^3 \sqrt{q} \log^3 n. \quad (6.1)$$

Before proving Theorem 6.1, we show how it implies Theorem 2.1.

Proof of Theorem 2.1 assuming Theorem 6.1. This is an elementary exercise, and we only sketch the argument. Part (1) of Theorem 2.1 is proved by dealing separately with the case $\xi < 1$ and $\xi \geq 1$. In both cases one can show that the assumption $nq \geq \frac{C(1+D)^4}{\min\{\xi^6, 1\}} \log^4 n$ implies that $\mathcal{E}(\xi/2) < n^{-D}$. Finally, note that under the same assumption we have $pd \leq \xi/(2\sqrt{nq})$ which allows us to use the interval $\sqrt{dnq}[-2-\xi, 2+\xi]$ instead of $-pd + \sqrt{dnq}[-2-\xi/2, 2+\xi/2]$. As for part (2) of Theorem 2.1, note that the assumption $nq \geq C(1+D)^4 \log^4 n$ implies that $\mathcal{E}((\sqrt{6}-2)\sqrt{d}) < n^{-D}/2$ for an appropriate choice of C . By choosing $\xi = C'(D+1)$ with C' depending only on d one can verify that $\mathcal{E}(C'(D+1))$ is bounded by $n^{-D}/2$ as well. Therefore as long as $nq \geq \frac{C'^2(D+1)^2 \log^6 n}{n}$ we have that with probability at least $1 - n^{-D}$ that the remaining eigenvalues are within the interval $np + [-\Gamma(C'(1+D), n), \Gamma(C'(1+D), n)]$ which under both assumption of part (2) is contained within the interval $np + [-7d, 7d]$ for an appropriate choice of the constant C . Finally, note that the case $nq \leq \frac{C'^2(D+1)^2 \log^6 n}{n}$ is impossible, since when combined with the assumption $nq \geq C(D+1)^4 \log^4 n$ implies that $(D+1)^2 \leq \frac{C'^2 \log^2 n}{C}$, which by increasing the value of the constant C does not hold for any value of n . \square

The rest of this section is devoted to the proof of Theorem 6.1. We introduce

$$P := \frac{1}{n}(\mathbb{A} + dI), \quad (6.2)$$

which by Lemma 4.1 is an orthogonal projection of rank $\binom{n-1}{d-1}$, and

$$\kappa := \frac{\sqrt{nq}}{1-p}. \quad (6.3)$$

Proof of Theorem 6.1 part (1). Recall from (4.1) that

$$(nq)^{-1/2}(A + pdI) = H + \kappa P, \quad (6.4)$$

Therefore, it follows from Theorem 5.1 that for every $\xi > 0$ with probability at least $1 - \mathcal{E}(\xi)$

$$\|(nq)^{-1/2}(A + pdI) - \kappa P\| = \|H\| \leq 2\sqrt{d} + \xi.$$

Since P is an orthogonal projection whose rank is $\binom{n-1}{d-1}$ it follows from Weyl's inequality that the eigenvalues of $(nq)^{-1/2}(A + pdI)$ are shifted by at most $2\sqrt{d} + \xi$ from the ones of κP , which in particular gives the required result for the $\binom{n-1}{d}$ smallest eigenvalues of A . \square

The above proof also implies that the remaining $\binom{n-1}{d-1}$ eigenvalues of A are within an interval of size \sqrt{nq} around κ with probability at least $1 - \mathcal{E}(\xi)$. Since we are interested in showing a better concentration result for those eigenvalues we need to obtain better estimation on the largest $\binom{n-1}{d-1}$ eigenvalues. More precisely, we have the following. Let $Y := (nq)^{-1/2}(A + pdI)$ and $\bar{P} := I - P$, and split

$$Y = \underbrace{(PYP + \bar{P}Y\bar{P})}_{Y_1} + \underbrace{(PY\bar{P} + \bar{P}YP)}_{Y_2}.$$

It follows that with probability at least $1 - \mathcal{E}(\xi)$ we have $\|Y_2\| \leq \|H\| \leq 2\sqrt{d} + \xi$ and that (due to (6.4)) the distance between the eigenvalues of $PYP = PHP + \kappa P$ and $\bar{P}Y\bar{P} = \bar{P}H\bar{P}$ is at least $\kappa - 2\|H\| \geq \kappa - 2(2\sqrt{d} + \xi)$. Consequently, by [KY14, Proposition A.1],

$$|\lambda_i(Y_1) - \lambda_i(Y)| \leq \frac{(2\sqrt{d} + \xi)^2}{\kappa - 4(2\sqrt{d} + \xi)} = O\left(\frac{1-p}{\sqrt{nq}}\right) \quad (6.5)$$

for every $\binom{n-1}{d} < i \leq \binom{n}{d}$. Since $Y_1 = PHP + \kappa P + \bar{P}H\bar{P}$ it follows that $\lambda_i(Y_1)$ for $\binom{n-1}{d} < i \leq \binom{n}{d}$ are shifted from κ by an amount which is at most the norm of the matrix PHP .

Combining all of the above we conclude that

$$\begin{aligned} |\lambda_i((nq)^{-1/2}A) - \kappa| &\leq |\lambda_i((nq)^{-1/2}A) - \lambda_i(Y)| + |\lambda_i(Y) - \lambda_i(Y_1)| + |\lambda_i(Y_1) - \kappa| \\ &\leq \frac{pd}{\sqrt{nq}} + \frac{(2\sqrt{d} + \xi)^2}{\kappa - 4(2\sqrt{d} + \xi)} + \|PHP\| \end{aligned} \quad (6.6)$$

for every $\binom{n-1}{d} < i \leq \binom{n}{d}$. Thus it is enough to obtain a better bound on the norm of PHP , which is the content of the following proposition.

Proposition 6.2. *For every $\eta > 0$, if $nq \geq \frac{\eta^2}{16d^6e^2} \frac{\log^6 n}{n}$, then*

$$\begin{aligned} \mathbb{P}\left(\|PHP\| > \eta \frac{\log^3 n}{\sqrt{n}}\right) \\ \leq \frac{4e^3 d^{5/2}}{(d-1)!} \exp\left(\frac{5}{3} \log\left(\frac{\eta}{4d^{7/2}e}\right) + 5 \log \log n - \left[\frac{1}{2} \left(\frac{\eta}{4d^{7/2}e}\right)^{1/3} - d\right] \log n\right). \end{aligned} \quad (6.7)$$

Before proving Proposition 6.2, we use it to conclude the proof of Theorem 6.1.

Proof of Theorem 6.1 part (2). It follows from Proposition 6.2 (with $\eta = 4d^{7/2}e(2d + 2\xi)^3$) that with probability at least $1 - \mathcal{E}(\xi)$ we have $\|PHP\| \leq 4d^{7/2}e(2d + 2\xi)^3 n^{-1/2} \log^3 n$ assuming that $nq \geq d(2d + 2\xi)^6 n^{-1} \log^6 n$. When combined with (6.6) (with the value $\xi = (\sqrt{6} - 2)\sqrt{d}$ in Theorem 6.1 part (1)) this yields that with probability at least $1 - \mathcal{E}((\sqrt{6} - 2)\sqrt{d}) - \mathcal{E}(\xi)$ we have $|\lambda_i(A) - np| \leq \Gamma(\xi, n)$ for every $\binom{n-1}{d} < i \leq \binom{n}{d}$, assuming that $nq \geq d(2d + 2\xi)^6 n^{-1} \log^6 n$. \square

6.1 Proof of Proposition 6.2

The proof of Proposition 6.2 is based on a variant of the FK method developed in Section 5. There are several differences between the proof for the matrix H and for the matrix PHP . First, since the matrix P is deterministic, the requirement that each d -cell must be crossed twice is only valid for crossings associated with entries of H , resulting in a weaker constraint on the set of admissible words. Second, in the Füredi-Komlós-type bound on the number of equivalence classes of words (see Lemma 6.7) we only obtain a rough bound, by showing that the words in the FK parsing contain at least $s - d$ crossings, instead of the $2(s - d)$ crossings proved in Lemma 5.10; we note that this bound may be improved using a more refined analysis, but it is sufficient for our purposes and has the advantage of having a relatively concise proof. Finally, the diagonal entries of the matrix P are non-zero, which leads to a slightly larger class of words. The losses resulting from the weaker restriction on the set of admissible words and the weaker lower bound on the number of crossings in words of the FK parsing are compensated by the additional factor n^{-1} associated with each entry of P , which ultimately allows us to improve the final estimate by the required factor $(nq)^{-1/2}$.

The proof starts with the following observation. Let $Q := nP = \mathbb{A} + dI$ and $B := A - \mathbb{E}[A]$. Then, analogously to (3.5), we have

$$\mathbb{E}\left[\int_{\mathbb{R}} x^k L_{PHP}(dx)\right] \leq \frac{1}{n^{d-1}(np)^{k/2}n^k} \sum_{\sigma_1, \dots, \sigma_{2k} \in X_+^{d-1}} |\mathbb{E}[B_{\sigma_1\sigma_2} Q_{\sigma_2\sigma_3} \dots B_{\sigma_{2k-1}\sigma_{2k}} Q_{\sigma_{2k}\sigma_1}]| \quad (6.8)$$

We wish to rewrite the last sum using closed word, similarly to (3.10). In order to do that a new definition for words is needed.

Definition 6.3 (2-words). An (n, d) -word of type 2 (or shortly a 2-word) is a finite sequence $\sigma_1 \dots \sigma_{2k} \sigma_{2k+1}$ of letters at least three letters long such that:

- $\sigma_{2i-1} \cup \sigma_{2i}$ is a d -cell for every $1 \leq i \leq k$.
- For every $1 \leq i \leq k$ either $\sigma_{2i} \cup \sigma_{2i+1}$ is a d -cell, or $\sigma_{2i} = \sigma_{2i+1}$.

The *length* of $w = \sigma_1 \dots \sigma_{2k+1}$ is defined to be $2k + 1$. A 2-word is called *closed* if its first and last letters are the same. Two 2-words $w = \sigma_1 \dots \sigma_{2k+1}$ and $w' = \sigma'_1 \dots \sigma'_{2k+1}$ are called *equivalent* if there exists a permutation π on $X^0 = V$ such that $\pi(\sigma_i) = \sigma'_i$ for every $1 \leq i \leq 2k + 1$ (see Definition 3.4 for the meaning of $\pi(\sigma_i)$).

Definition 6.4 (Support of 2-words). For a 2-word $w = \sigma_1 \dots \sigma_{2k+1}$ we define its *support* $\text{supp}_0(w) = \sigma_1 \cup \sigma_2 \cup \dots \cup \sigma_{2k+1}$, its *d -cell support*

$$\text{supp}_d(w) = \{\sigma_i \cup \sigma_{i+1} : 1 \leq i \leq 2k \text{ such that } |\sigma_i \cup \sigma_{i+1}| = d + 1\},$$

and its *odd-crossing d -cell support* $\text{supp}_d^{\text{odd}}(w) = \{\sigma_{2j-1} \cup \sigma_{2j} : 1 \leq j \leq k\}$.

Definition 6.5 (The graph of a 2-word). Given a 2-word $w = \sigma_1 \dots \sigma_{2k+1}$ we define $G_w = (V_w, E_w)$ to be the graph with vertex set $V_w = \{\sigma_i : 1 \leq i \leq 2k + 1\}$ and edge set $E_w = \{\{\sigma_i, \sigma_{i+1}\} : 1 \leq i \leq 2k\}$; note that we may have $\sigma_i = \sigma_{i+1}$, in which case $\{\sigma_i, \sigma_{i+1}\}$ gives rise to a loop in E_w . For an edge $e \in E_w$ we define $N_w(e) = |\{1 \leq i \leq 2k + 1 : \{\sigma_i, \sigma_{i+1}\} = e\}|$ to be the number of times the edge e is crossed along the path generated by w in the graph G_w , and let $N_w^{\text{odd}}(e) = |\{1 \leq j \leq k : \{\sigma_{2j-1}, \sigma_{2j}\} = e\}|$ be the number of crossings made in odd steps. As in the case of words we define the i -th crossing time of an edge and of a d -cell. In addition, the edges of the graph G_w that are not loops can be divided into different classes according to the d -cell generated by the two $(d - 1)$ -cells which form its vertices. For a d -cell τ we define $E_w(\tau) = \{\{\sigma, \sigma'\} \in E_w : \sigma \cup \sigma' = \tau\}$ and $N_w(\tau) = \sum_{e \in E_w(\tau)} N_w(e)$, and let $N_w^{\text{odd}}(\tau) = \sum_{e \in E_w(\tau)} N_w^{\text{odd}}(e)$.

Using the above definitions and going back to (6.8) we see that each term in the sum can be associated with a list of letters $w = \sigma_1 \sigma_2 \dots \sigma_{2k} \sigma_{2k+1}$ with $\sigma_{2k+1} = \sigma_1$. Since $B_{\sigma, \sigma'} = 0$ whenever $\sigma \cup \sigma' \notin X^d$ and $Q_{\sigma, \sigma'} = 0$ unless $\sigma \cup \sigma' \in X^d$ or $\sigma = \sigma'$ it follows that we can restrict the sum in (6.8) to those $w = \sigma_1 \dots \sigma_{2k+1}$ which are closed 2-words of length $2k + 1$. Using the independence structure of A for different d -cells and the definitions of N_w and N_w^{odd} we then have

$$\begin{aligned} \mathbb{E} \left[\int_{\mathbb{R}} x^k L_{PHP}(dx) \right] &\leq \frac{1}{n^{d-1} (nq)^{\frac{k}{2}} n^k} \sum_{\substack{w \text{ a closed 2-word} \\ \text{of length } k+1}} \left| \prod_{\tau \in X^d} \mathbb{E} \left[(\chi - p)^{N_w^{\text{odd}}(\tau)} \right] \prod_{\sigma \in X^{d-1}} d^{N_w(\{\sigma, \sigma\})} \right| \\ &\leq \frac{d^{2k+1}}{n^{d-1} (nq)^{\frac{k}{2}} n^k} \sum_{\substack{w \text{ a closed 2-word} \\ \text{of length } k+1}} \left| \prod_{\tau \in X^d} \mathbb{E} \left[(\chi - p)^{N_w^{\text{odd}}(\tau)} \right] \right|. \end{aligned}$$

Since $\mathbb{E}[\chi - p] = 0$, it follows that we can restrict the last sum to those 2-words for which $N_w^{\text{odd}}(\tau) \neq 1$ for every $\tau \in X^d$.

Next, we rewrite the sum over such words using their equivalence classes. Denoting by $[w]$ the equivalence class of a 2-word w , we have

$$\mathbb{E} \left[\int_{\mathbb{R}} x^k L_{PHP}(dx) \right] \leq \frac{d^{2k+1}}{n^{d-1} (nq)^{\frac{k}{2}} n^k} \sum_w |[w]| \left| \prod_{\tau \in X^d} \mathbb{E} \left[(\chi - p)^{N_w^{\text{odd}}(\tau)} \right] \right|, \quad (6.9)$$

where the sum is over representatives of the equivalence classes of closed 2-words of length $2k+1$ such that $N_w^{\text{odd}}(\tau) \neq 1$ for every $\tau \in X^d$.

We further distinguish between different equivalence classes according to the number of 0-cells in $\text{supp}_0(w)$. Denoting by $\widehat{\mathcal{W}}_s^k = \widehat{\mathcal{W}}_s^k(n, d)$ a set of representatives for the equivalence classes of closed 2-words of length $2k+1$ such that $N_w^{\text{odd}}(\tau) \neq 1$ for all $\tau \in X^d$ and $|\text{supp}_0(w)| = s$, and observing that by the same argument used to prove Claim 3.9 the number of elements in the equivalence class of $w \in \widehat{\mathcal{W}}_s^k$ is $B_{n,s,d} \equiv \frac{n(n-1)\cdots(n-s)}{d!}$, we can rewrite (6.9) as

$$\begin{aligned} \mathbb{E} \left[\int_{\mathbb{R}} x^k L_{PHP}(\mathrm{d}x) \right] &\leq \sum_{s \geq d} \sum_{w \in \widehat{\mathcal{W}}_s^k} \frac{d^{2k+1} n^s}{n^{d-1} (nq)^{\frac{k}{2}} n^k} \left| \prod_{\tau \in X^d} \mathbb{E} \left[(\chi - p)^{N_w^{\text{odd}}(\tau)} \right] \right| \\ &\leq \sum_{s \geq d} \frac{d^{2k+1} n^s q^{|\text{supp}_d^{\text{odd}}(w)|}}{n^{d-1} (nq)^{\frac{k}{2}} n^k} |\widehat{\mathcal{W}}_s^k|, \end{aligned} \quad (6.10)$$

where for the second inequality we used the same argument that yields (3.12).

The next lemma gives constraints on the value of $|\text{supp}_d^{\text{odd}}(w)|$ for a given s as well as a bound on the values of s itself.

Lemma 6.6. *We have $\frac{|\text{supp}_0(w)|-d}{2} \leq |\text{supp}_d^{\text{odd}}(w)| \leq \lfloor \frac{k}{2} \rfloor$, and therefore $|\text{supp}_0(w)| \leq k+d$.*

Proof. Recalling that $N_{\text{odd}}(\tau) \neq 1$ for all $\tau \in X^d$ and using the fact that $\sum_{\tau \in X^d} N_{w,\text{odd}}(\tau) = k$ it follows that $|\text{supp}_d^{\text{odd}}(w)| \leq \lfloor \frac{k}{2} \rfloor$. As for the lower bound, note that if σ_j is the first appearance of a new 0-cell, then both $\sigma_{i-1}\sigma_i$ and $\sigma_i\sigma_{i+1}$ are crossings in which this 0-cell appeared. Consequently, the d -cells crossed in odd times contain all 0-cells in $\text{supp}_0(w)$, and each odd crossing can reveal at most two new 0-cells. \square

Combining (6.10) and Lemma 6.6 yields

$$\mathbb{E} \left[\int_{\mathbb{R}} x^k L_{PHP}(\mathrm{d}x) \right] \leq \frac{d^{2k+1}}{(nq)^{\frac{k}{2}} n^{k-1}} \sum_{s=d}^{k+d} (n\sqrt{q})^{s-d} |\widehat{\mathcal{W}}_s^k|. \quad (6.11)$$

What remains is an Füredi-Komlós-type bound on $|\widehat{\mathcal{W}}_s^k|$, which is the content of the following lemma.

Lemma 6.7. *For every $d \leq s \leq k+d$*

$$|\widehat{\mathcal{W}}_s^k| \leq 2d(4d)^k (2k+1)(\sqrt{d}(2k+1)^3)^{2k+1-(s-d)}$$

Before proving Lemma 6.7 we use it to conclude the proof of Proposition 6.2. Combining (6.11) and Lemma 6.7 we get for even k

$$\begin{aligned} \mathbb{E} \left[\int_{\mathbb{R}} x^k L_{PHP}(\mathrm{d}x) \right] &\leq 2d^2(2k+1) \frac{(4d^3)^k (\sqrt{d}(2k+1)^3)^{2k+1}}{(nq)^{\frac{k}{2}} n^{k-1}} \sum_{s=d}^{k+d} \left(\frac{n\sqrt{q}}{\sqrt{d}(2k+1)^3} \right)^{s-d} \\ &\leq 2d^2(2k+1) \frac{(4d^3)^k (\sqrt{d}(2k+1)^3)^{2k+1}}{(nq)^{\frac{k}{2}} n^{k-1}} (k+1) \left[\left(\frac{n\sqrt{q}}{\sqrt{d}(2k+1)^3} \right)^k + 1 \right] \\ &\leq 2d^{5/2}(2k+1)^5 n \left[\left(\frac{4d^{7/2}(2k+1)^3}{\sqrt{n}} \right)^k + \left(\frac{4d^4(2k+1)^6}{\sqrt{nqn}} \right)^k \right]. \end{aligned}$$

It follows from Markov's inequality that for every even k

$$\begin{aligned} \mathbb{P} \left(\|PHP\| > \eta \frac{\log^3 n}{\sqrt{n}} \right) \\ \leq 2d^{5/2}(2k+1)^5 \frac{n^d}{(d-1)!} \left[\left(\frac{4d^{7/2}(2k+1)^3}{\eta \log^3 n} \right)^k + \left(\frac{4d^4(2k+1)^6}{\sqrt{q} n \eta \log^3 n} \right)^k \right]. \end{aligned}$$

By taking $k = k(n)$ to be the largest even integer which is smaller than $\frac{1}{2}(\eta/(4d^{7/2}e))^{1/3} \log n - 1$ and assuming that $nq \geq \frac{\eta^2}{16d^6e^2} \frac{\log^6 n}{n}$ we obtain that

$$\begin{aligned} \mathbb{P} \left(\|PHP\| > \eta \frac{\log^3 n}{\sqrt{n}} \right) \\ \leq \frac{4e^3 d^{5/2}}{(d-1)!} \exp \left(\frac{5}{3} \log \left(\frac{\eta}{4d^{7/2}e} \right) + 5 \log \log n - \left[\frac{1}{2} \left(\frac{\eta}{4d^{7/2}e} \right)^{1/3} - d \right] \log n \right), \end{aligned}$$

thus completing the proof. \square

Proof of Lemma 6.7. We closely follow the proof of Proposition 5.3. Given $w \in \widehat{\mathcal{W}}_s^k$ we define its FK parsing a_w by parsing w in the crossing $\sigma_i \sigma_{i+1}$ if one of the following occurs:

- $\{\sigma_i, \sigma_{i+1}\}$ is an old edge of G_w .
- $\{\sigma_i, \sigma_{i+1}\}$ is a third or subsequent crossing of the edge $\{\sigma_i, \sigma_{i+1}\}$.
- $\{\sigma_i, \sigma_{i+1}\}$ is a second crossing of the edge $\{\sigma_i, \sigma_{i+1}\}$ and there exists an edge $\{\sigma, \sigma'\} \in E_w(\sigma_i \cup \sigma_{i+1})$ in the d -cell as $\sigma_i \cup \sigma_{i+1}$ which is crossed twice in $\sigma_1 \sigma_2 \dots \sigma_i$.
- $\{\sigma_i, \sigma_{i+1}\}$ is a loop.

The resulting sentence is an FK sentence in the sense of Definition 5.8. Since the path of each word in $\widehat{\mathcal{W}}_s^k$ must contain $s - d$ crossing times in which a new 0-cell is observed, it follows that the number of words m_w in the FK sentence a_w satisfies

$$1 \leq m_w \leq 2k + 1 - (s - d).$$

By repeating the same argument as in the proof of Proposition 5.3 and using Lemma 5.11 and Lemma 5.12 we obtain

$$\begin{aligned} |\widehat{\mathcal{W}}_s^k| &\leq \sum_{m=1}^{2k+1-(s-d)} \binom{2k}{m-1} \left(\frac{\sqrt{d}}{2} \right)^m (2\sqrt{d})^{2k+1} (2k+1)^{2(m-1)} \\ &\leq 2d(4d)^k (2k+1)(\sqrt{d}(2k+1)^3)^{2k+1-(s-d)}, \end{aligned}$$

thus completing the proof. \square

References

- [AGZ10] Greg W. Anderson, Alice Guionnet, and Ofer Zeitouni. *An introduction to random matrices*, volume 118 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.
- [ALLM13] Lior Aronshtam, Nathan Linial, Tomasz Łuczak, and Roy Meshulam. Collapsibility and vanishing of top homology in random simplicial complexes. *Discrete Comput. Geom.*, 49(2):317–334, 2013.

- [Alo86] Noga Alon. Eigenvalues and expanders. *Combinatorica*, 6(2):83–96, 1986. Theory of computing (Singer Island, Fla., 1984).
- [AM85] Noga Alon and Vitali D. Milman. λ_1 , isoperimetric inequalities for graphs, and superconcentrators. *J. Combin. Theory Ser. B*, 38(1):73–88, 1985.
- [BHK11] Eric Babson, Christopher Hoffman, and Matthew Kahle. The fundamental group of random 2-complexes. *J. Amer. Math. Soc.*, 24(1):1–28, 2011.
- [CCFK12] Daniel Cohen, Armindo Costa, Michael Farber, and Thomas Kappeler. Topology of random 2-complexes. *Discrete Comput. Geom.*, 47(1):117–149, 2012.
- [Chu97] Fan R. K. Chung. *Spectral graph theory*, volume 92 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC; by the American Mathematical Society, Providence, RI, 1997.
- [CMRT14] Emma Cohen, Dhruv Mubayi, Peter Ralli, and Prasad Tetali. Inverse expander mixing for hypergraphs. *arXiv preprint arXiv:1407.2285*, 2014.
- [CO07] Amin Coja-Oghlan. On the Laplacian eigenvalues of $G_{n,p}$. *Combin. Probab. Comput.*, 16(6):923–946, 2007.
- [CRS10] Dragoš Cvetković, Peter Rowlinson, and Slobodan Simić. *An introduction to the theory of graph spectra*, volume 75 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 2010.
- [CSŽ03] Donald I. Cartwright, Patrick Solé, and Andrzej Żuk. Ramanujan geometries of type A_n . *Discrete Math.*, 269(1-3):35–43, 2003.
- [DK12] Dominic Dotterrer and Matthew Kahle. Coboundary expanders. *J. Topol. Anal.*, 4(4):499–514, 2012.
- [DKW15] Dominic Dotterrer, Tali Kaufman, and Uli Wagner. On expansion and topological overlap. *arXiv preprint arXiv:1506.04558*, 2015.
- [Eck45] Beno Eckmann. Harmonische Funktionen und Randwertaufgaben in einem Komplex. *Comment. Math. Helv.*, 17:240–255, 1945.
- [EGL15] Shai Evra, Konstantin Golubev, and Alexander Lubotzky. Mixing properties and the chromatic number of Ramanujan complexes. *International Mathematics Research Notices*, pages 1–29, 2015.
- [EKYY12] László Erdős, Antti Knowles, Horng-Tzer Yau, and Jun Yin. Spectral statistics of Erdős-Rényi Graphs II: Eigenvalue spacing and the extreme eigenvalues. *Comm. Math. Phys.*, 314(3):587–640, 2012.
- [EKYY13] László Erdős, Antti Knowles, Horng-Tzer Yau, and Jun Yin. Spectral statistics of Erdős-Rényi graphs I: Local semicircle law. *Ann. Probab.*, 41(3B):2279–2375, 2013.
- [ER59] Paul Erdős and Alfréd Rényi. On random graphs. I. *Publ. Math. Debrecen*, 6:290–297, 1959.
- [ER61] Paul Erdős and Alfréd Rényi. On the evolution of random graphs. *Bull. Inst. Internat. Statist.*, 38:343–347, 1961.
- [Evr15] Shai Evra. Finite quotients of Bruhat-Tits buildings as geometric expanders. *arXiv preprint arXiv:1503.08626*, 2015.
- [FGL⁺12] Jacob Fox, Mikhail Gromov, Vincent Lafforgue, Assaf Naor, and János Pach. Overlap properties of geometric expanders. *J. Reine Angew. Math.*, 671:49–83, 2012.
- [FK81] Zoltán Füredi and János Komlós. The eigenvalues of random symmetric matrices. *Combinatorica*, 1(3):233–241, 1981.
- [FKS89] Joel Friedman, Jeff Kahn, and Endre Szemerédi. On the second eigenvalue of random regular graphs. In *Proceedings of the twenty-first annual ACM symposium on Theory of computing*, pages 587–598. ACM, 1989.

- [FO05] Uriel Feige and Eran Ofek. Spectral techniques applied to sparse random graphs. *Random Structures Algorithms*, 27(2):251–275, 2005.
- [Gar73] Howard Garland. p -adic curvature and the cohomology of discrete subgroups of p -adic groups. *Ann. of Math. (2)*, 97:375–423, 1973.
- [Gol13] Konstantin Golubev. On the chromatic number of a simplicial complex. *arXiv preprint arXiv:1306.4818*, 2013.
- [GP14] Konstantin Golubev and Ori Parzanchevski. Spectrum and combinatorics of Ramanujan triangle complexes. *arXiv preprint arXiv:1406.6666*, 2014.
- [Gro10] Mikhail Gromov. Singularities, expanders and topology of maps. Part 2: From combinatorics to topology via algebraic isoperimetry. *Geom. Funct. Anal.*, 20(2):416–526, 2010.
- [GS14] Anna Gundert and May Szedlák. Higher dimensional Cheeger inequalities. In *Annual Symposium on Computational Geometry*, page 181. ACM, 2014.
- [GW14] Anna Gundert and Uli Wagner. On eigenvalues of random complexes. *arXiv preprint arXiv:1411.4906*, 2014.
- [Hat02] Allen Hatcher. *Algebraic topology*. Cambridge University Press, Cambridge, 2002.
- [HJ13] Danijela Horak and Jürgen Jost. Spectra of combinatorial Laplace operators on simplicial complexes. *Adv. Math.*, 244:303–336, 2013.
- [HKP12] Christopher Hoffman, Matthew Kahle, and Elliot Paquette. Spectral gaps of random graphs and applications to random topology. *arXiv preprint arXiv:1201.0425*, 2012.
- [HKP13] Christopher Hoffman, Matthew Kahle, and Elliot Paquette. The threshold for integer homology in random d -complexes. *arXiv preprint arXiv:1308.6232*, 2013.
- [HLW06] Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bull. Amer. Math. Soc. (N.S.)*, 43(4):439–561 (electronic), 2006.
- [KKL14] Tali Kaufman, David Kazhdan, and Alexander Lubotzky. Isoperimetric inequalities for Ramanujan complexes and topological expanders. *arXiv preprint arXiv:1409.1397*, 2014.
- [Koz10] Dmitry N. Kozlov. The threshold function for vanishing of the top homology group of random d -complexes. *Proc. Amer. Math. Soc.*, 138(12):4517–4527, 2010.
- [KS06] Michael Krivelevich and Benny Sudakov. Pseudo-random graphs. In *More sets, graphs and numbers*, volume 15 of *Bolyai Soc. Math. Stud.*, pages 199–262. Springer, Berlin, 2006.
- [KY14] Antti Knowles and Jun Yin. The outliers of a deformed Wigner matrix. *Ann. Probab.*, 42(5):1980–2031, 2014.
- [Li04] Wen-Ching W. Li. Ramanujan hypergraphs. *Geom. Funct. Anal.*, 14(2):380–399, 2004.
- [LM06] Nathan Linial and Roy Meshulam. Homological connectivity of random 2-complexes. *Combinatorica*, 26(4):475–487, 2006.
- [LM15] Alexander Lubotzky and Roy Meshulam. Random Latin squares and 2-dimensional expanders. *Adv. Math.*, 272:743–760, 2015.
- [LP14] Nathan Linial and Yuval Peled. On the phase transition in random simplicial complexes. *arXiv preprint arXiv:1410.1281*, 2014.
- [LSV05] Alexander Lubotzky, Beth Samuels, and Uzi Vishne. Ramanujan complexes of type A_d . *Israel J. Math.*, 149:267–299, 2005. Probability in mathematics.
- [Lub14] Alexander Lubotzky. Ramanujan complexes and high dimensional expanders. *Jpn. J. Math.*, 9(2):137–169, 2014.
- [MS13] Sayan Mukherjee and John Steenbergen. Random walks on simplicial complexes and harmonics. *arXiv preprint arXiv:1310.5099*, 2013.

- [MW09] Roy Meshulam and Nathan Wallach. Homological connectivity of random k -dimensional complexes. *Random Structures Algorithms*, 34(3):408–417, 2009.
- [MW14] Jiří Matoušek and Uli Wagner. On Gromov’s method of selecting heavily covered points. *Discrete Comput. Geom.*, 52(1):1–33, 2014.
- [Nil91] Alon Nilli. On the second eigenvalue of a graph. *Discrete Math.*, 91(2):207–210, 1991.
- [Opp14] Izhar Oppenheim. Local spectral expansion approach to high dimensional expanders. *arXiv preprint arXiv:1407.8517*, 2014.
- [Par13] Ori Parzanchevski. Mixing in high-dimensional expanders. *arXiv preprint arXiv:1310.6477*, 2013.
- [PR12] Ori Parzanchevski and Ron Rosenthal. Simplicial complexes: spectrum, homology and random walks. *arXiv preprint arXiv:1211.6775*, 2012.
- [PRT15] Ori Parzanchevski, Ron Rosenthal, and Ran J. Tessler. Isoperimetric inequalities in simplicial complexes. *Combinatorica*, pages 1–33, 2015.
- [Ros14] Ron Rosenthal. Simplicial branching random walks and their applications. *arXiv preprint arXiv:1412.5406*, 2014.
- [SKM14] John Steenbergen, Caroline Klivans, and Sayan Mukherjee. A Cheeger-type inequality on simplicial complexes. *Adv. in Appl. Math.*, 56:56–77, 2014.
- [Wag11] Uli Wagner. Minors in random and expanding hypergraphs. In *Computational geometry (SCG’11)*, pages 351–360. ACM, New York, 2011.
- [Wig58] Eugene P. Wigner. On the distribution of the roots of certain symmetric matrices. *Ann. of Math. (2)*, 67:325–327, 1958.

Antti Knowles

E-mail: knowles@math.ethz.ch

Ron Rosenthal

E-mail: ron.rosenthal@math.ethz.ch

Department Mathematik

ETH Zürich

CH-8092 Zürich

Switzerland